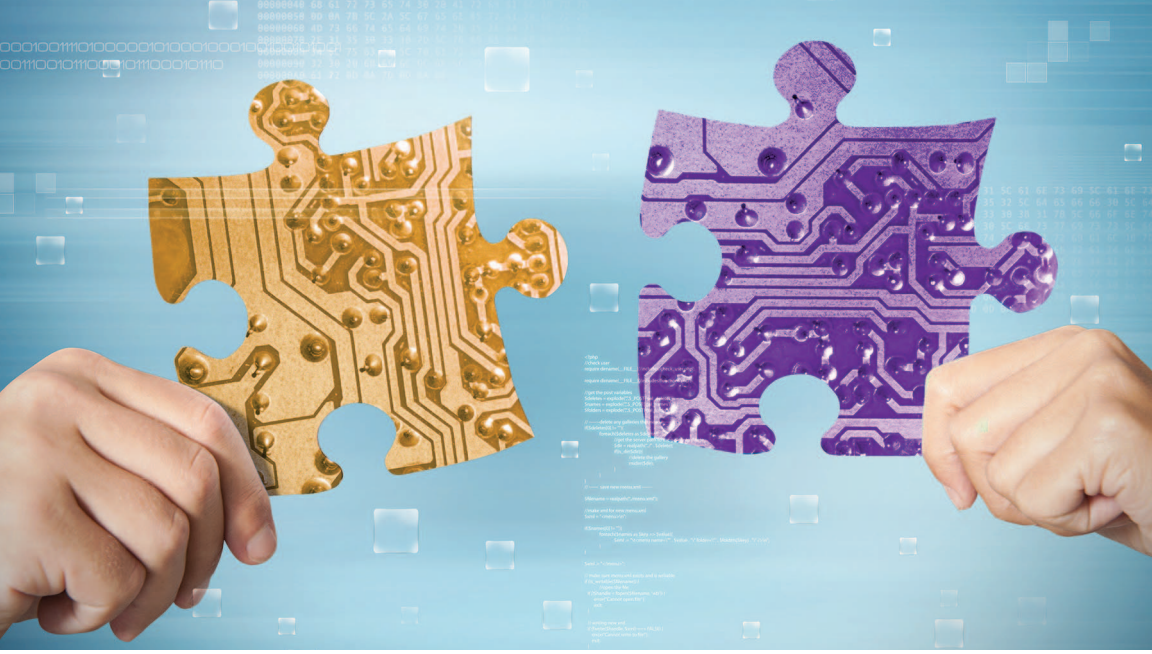


Uncertainty-Sensitive Heterogeneous Information Fusion

Assessing Threat with Soft, Uncertain, and Conflicting Evidence

Paul K. Davis, Walter L. Perry, John S. Hollywood, David Manheim

RAND CORPORATION

For more information on this publication, visit www.rand.org/t/RR1200

Library of Congress Control Number: 2016931460
ISBN: 978-0-8330-9277-9

Published by the RAND Corporation, Santa Monica, Calif.

Cover Image: Fotolia

Preface

This study was sponsored by the Office of Naval Research (ONR). It was initiated by Ivy Estabrooke and continued under Lee Mastrioanni, thrust manager, and Joong Kim, program officer, in ONR's Expeditionary Maneuver Warfare and Combatting Terrorism Department (Code 30). Comments and questions are welcome and should be addressed to Paul K. Davis at pdavis@rand.org.

The report documents a basic research project. It is technical in nature and intended for researchers or managers of technical research potentially interested in information fusion for such domains as counterterrorism, law enforcement, and intelligence. Some of the ideas and methods will be of interest to the larger community of researchers involved with information fusion.

This research was sponsored by the Office of Naval Research and conducted within the International Security and Defense Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community.

For more information on the International Security and Defense Policy Center, see http://www.rand.org/nsrd/ndri/centers/isdp.html or contact the director (contact information is provided on web page).

Contents

Figures

Tables

Summary

Objectives

Military and other government organizations put substantial effort into detecting and thwarting attacks such as those by suicide bombers or involving improvised explosive devices. Such attacks may be against military or government installations in the United States or abroad, civilian infrastructure, or any of many other targets. An element of thwarting such attacks is observing suspicious individuals over time with such diverse means as cameras, scanners, and other devices; travel records; behavioral observations; and intelligence sources. Such observations provide data that are often both complex and "soft"—i.e., qualitative, subjective, fuzzy, or ambiguous—and also contradictory or even deceptive (as when humans lie). The problem, then, is how to fuse the heterogeneous data. This report summarizes our research on heterogeneous fusion methods. The context is military and civilian counterterrorism, but the ideas apply more generally in intelligence and law enforcement.

The ultimate objectives, which go well beyond what our project sought to accomplish, include improving real-world ability to detect terrorists, reducing the rate of false alarms, and making it easier to exonerate those who inappropriately come under suspicion. Previous RAND research discussed these and recommended investigation of heterogeneous information fusion that would be analyst-centric with flexible man-machine investigation to supplement more automated and data-driven activities. The research, then, had the objective of designing, implementing, and testing a research prototype with the *potential* for contributing to the ultimate objectives. The effort would illustrate

an integrative vision and assess technical feasibility. It would lay the foundations for the considerable subsequent work that would be necessary to assess real-world value and practicality.

Strategy and Implied Tasks

Rather early, we decided on a strategy to seek certain specific attributes in our prototype system. It should:

1. Address matters probabilistically rather than seeking point conclusions.
2. Routinely use a mix of methods for information fusion.
3. Allow for parallel, competitive, and iterative streams of analysis.
4. Employ cause-effect models (causal models).
5. Routinely explore consequences of uncertainties affecting models, process, and assumptions.

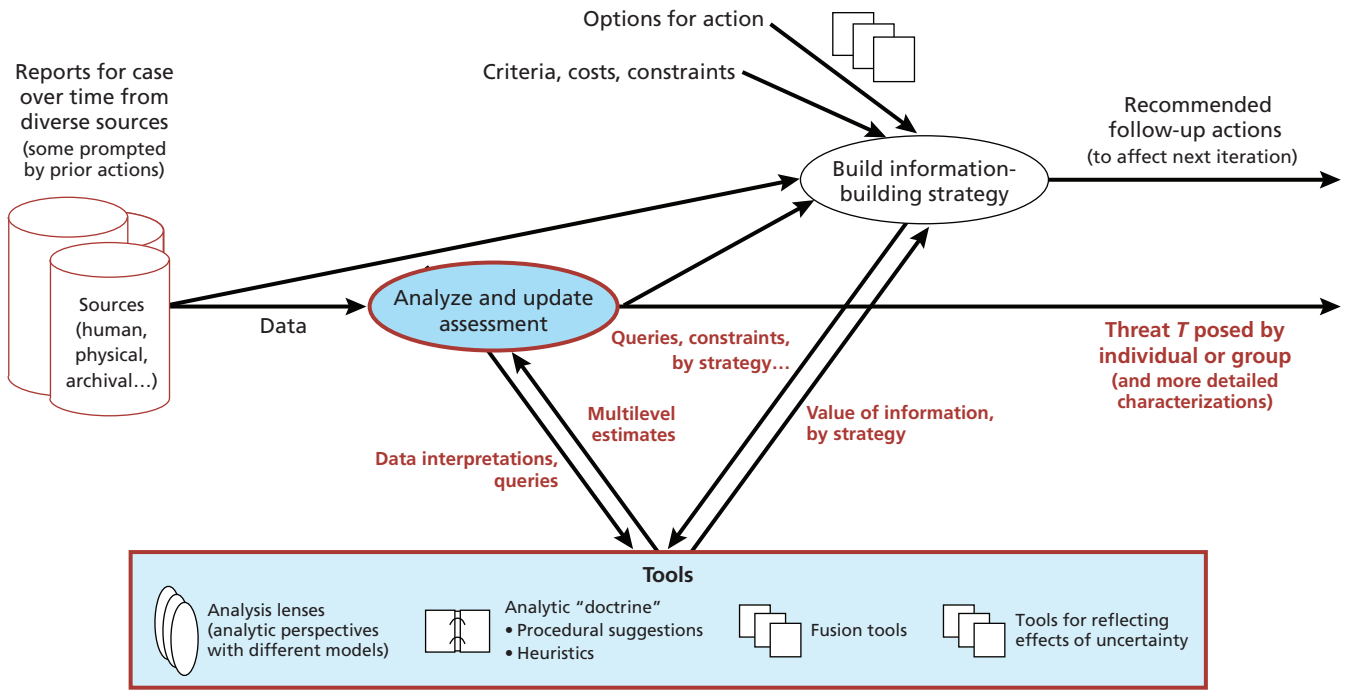
The system should, for those with technical backgrounds, be comprehensible, controllable, and analyst-friendly.

This strategy implied challenging research tasks: (1) designing an analytic architecture, (2) representing heterogeneous information effectively in probabilistic terms, (3) identifying suitable causal models for the prototype work, (4) adapting or developing appropriately different fusion methods, (5) developing synthetic test data to challenge the prototype systematically, and (6) experimentation. The report describes this research.

Architecture

Figure S.1 indicates the idealized vision that framed our work. For a given case (i.e., assessing the threat posed by a particular individual), information would derive from various sources, as shown on the far left. As information becomes available, information would be fused, drawing on an entire suite of methods and tools (bottom). This analysis

Figure S.1
Idealized Fusion



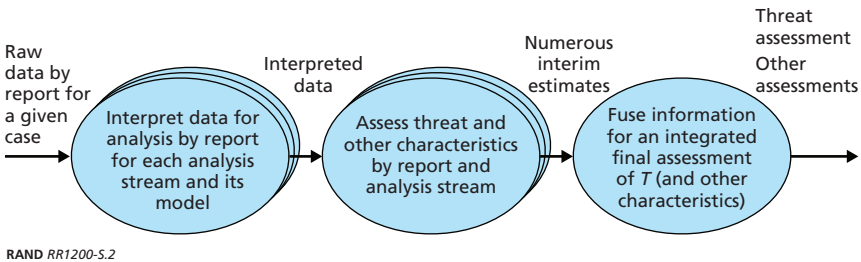
RAND *RR1200-S.1*

would update an assessment of the threat posed by the subject, painting an explanatory story where possible. Also, it would identify information that, if obtained, would be especially useful in further sharpening the assessment. This could inform follow-up decisions on information-gathering and other actions (top right of figure). Our research for this project focused on the analysis-related activities indicated by red letters or borders.

Figure S.2 describes the architecture of the prototype system, with the stacking of ovals indicating multiple streams of analysis. Each stream could use its own causal model and make its own choices about which methods and tools from Figure S.1 to employ. The streams might end up relatively similar or quite different. Information fusion would be conducted within and across streams. Cross-stream fusion might be a last step or might occur at several points. The analysis output would include an updated probabilistic threat assessment, e.g., characterizations of the individual's motivation or capability for terrorism. The process in Figure S.2 would continue over time and, *with enough information*, could yield a rich assessment of the individual's relationship to terrorism, if any.

Although apparently simple, the architecture implied by Figure S.2 had to be highly flexible to allow different intra-stream and inter-stream information fusion and to examine myriad alternative assumptions and choices. The assumptions and choices should be explicit to allow review, debate, and revision. Overall, the prototype reflected a

Figure S.2
Initial View of Heterogeneous Fusion Process



RAND *RR1200-S.2*

network-centric design rather than a classic procedural process specifying rigidly who does what with which tools, in what order.

Representing Information

Representing uncertainty-laden information is difficult for many reasons. Table S.1 summarizes challenges (left column) and how we came to deal with them. Moving downward through the table, the first block of rows deals with uncertainties. The first row distinguishes between uncertainties of knowledge and uncertainties due to random processes. The second notes the problem of indirect and fragmentary information, which is mitigated by using causal threat-assessment models. Dealing with correlations (third row) is a general and difficult challenge, which we dealt with by designing the threat-assessment model to have factors that are ordinarily independent. We also allowed, in exceptional cases, for representing correlations and testing effects of such correlations on conclusions. Concern about correlations should also affect the way reports are clumped and weighted, and how uncertainty analysis is conducted (e.g., recognizing that sources of information might have common mental models or "stories" causing seemingly independent information to be correlated).

The second block of challenges relates to combining complex information expressed as and's, or's, not's, inequalities, or conditional probabilities. We used multidimensional probability distributions, mathematics that preserves disjunctions, *sets* of alternative generic Bayesian likelihood functions, and mathematical index representations of qualitative statements.

The last block of challenges deals with conflicts, disagreements, and deception. It was essential to distinguish between disagreements that should be highlighted in fusion mathematics and those that should be smoothed over by using convergence-inducing mathematics. In either case, it is necessary to make explicit subjective assessments of the quality and ultimate reliability of different information. For dealing with all of these, exploratory analysis of uncertainty is essential—viewing results as the many assumptions are changes simultaneously.

Table S.1
Challenges and Responses in Representing Information

Challenges	Example	Representations
	Dealing with Uncertainties	
Uncertain knowledge versus randomness	Information on a factor may be uncertain because of limitations in knowledge or, sometimes, because attribute is changing, perhaps randomly.	Probability distributions for characterizing uncertain deterministic knowledge and for reflecting random processes.
Indirect and fragmentary knowledge	Is he a threat? One report: "He seems motivated"; a second report: "He has capability." Neither report covers all the factors affecting threat.	Threat assessment models using separately observable factors and filling in missing information with, e.g., empirical base rates.
Correlations	Information that is seemingly independent may be highly correlated.	Models with independent observable factors. Mechanisms for representing correlations. Sensitivity testing.
	Combining Information	
Disjunction (or)	Report: "Motivation is very high *or* very low, but not both."	Bimodal distributions. Fusion methods that preserve disjunctions.
Conjunction (and)	Report: "Motivation is high and so is capability."	Multi-attribute (multi-factor) distributions.
Complement (not)	Report: "He is not a threat; he has no motivation."	Assign probabilities accordingly.
Inequality	Report: "Motivation is more than nothing but less than very high."	Algebraic inequalities connecting distributions, e.g., very low < motivation < very high.
Likelihoods (conditional probabilities)	Given previous information, how should new information be interpreted?	Approximate Bayesian likelihood functions, where applicable, but with uncertainty analysis using alternative functions.

Table S.1—Continued

Challenges	Example	Representations
	Dealing with Conflicts, Disagreements, and Deception	
Disagreement to be resolved	First source says sees definite motivation; a second report sees the absence of motivation.	Fusion methods that preserve disagreement. Quality assessment of reports.
Disagreement to be reflected or reduced	Sources give estimates of motivation that vary from very low to very high, with no pattern.	Fusion methods that average or that estimate underlying convergence.
Deception about knowledge	Sources disagree on motivation but some are fabricating their stories.	Distinct reports assessed subjectively for quality and reliability.
Deception with intent to cause mis-assessment	Sources disagree; some are deliberately misleading: personal enemies may plant false negative information and friends may plant false positive information.	Distinct reports with associated reliabilities. Fusion methods that preserve conflicts. Alternative stories (higher-level "explanations").

This can also identify which pieces of information are most important. This has implications for next steps, whether expending resources to obtain additional information or taking intrusive or even preemptive steps against an individual.

Some of the difficulties of Table S.1 reflect the context of counter-terrorism, intelligence, and law enforcement more broadly. Much of the information does not come from "experts" in the normal sense of that term; some of the sources may be fabricating or even misleading; some may be making inferences based on erroneous or deceptive data. Lest such possibilities be obscured, the result of information fusion should be multimodal probabilities. The information-fusion challenge, then, is rather different from aggregating expert forecasts.

Causal Models

As recognized in science generally and in Bayesian-network fusion research specifically, causal models (i.e., those that represent cause-effect relationships rather than mere statistical correlations) can be very

helpful for inference from fragmentary and messy data. Often, however, such causal models are themselves very uncertain, and alternative models must be considered. We illustrate this by having two alternatives in the prototype. The first is qualitatively rooted in previous social-science research dealing with motivation for a cause or activity, perceptions about the legitimacy of terrorist acts, capability and opportunity for such terrorist acts, and acceptance of the costs and risks of taking such acts. The second model focuses merely on capability and opportunity, with no attempt to characterize the individual's cognition.

Fusion Methods

After considering numerous fusion methods discussed in the literature, we focused on a smaller set: (1) purely subjective assessments as the de facto baseline, (2) several algebraic fusion methods, (3) a quasi-Bayesian method, and (4) one that we called the maximum entropy/minimum penalty (MEMP) method. Although building on prior methods, we did a good deal of combining, adapting, and extending. Some of these methods are used for fusing information about causal factors to estimate threat (something we refer to as "combining" in the main text), some for fusing information across reports, and some for both.

By purely subjective assessments, we had in mind assessments expressed directly in light of the available evidence without the benefit of explicit analytical methods (a common practice currently).

We used three algebraic methods: linear weighted sums (LWS), thresholded linear weighted sums (TLWS), and primary factors (PF). LWS has been used extensively in past work on aggregating expert judgment and machine learning. The TLWS variant is a simple way to represent knowledge such that an effect depends on *all* of several factors being present to a significant degree. The causal model that we used for illustrative purposes assesses the threat posed by an individual to be low unless the individual has motivation; a perception of terrorist acts being legitimate; capability and opportunity; *and* a willingness to accept costs and risks. The PF method is an alternative chosen to be complementary. Subject to some constraints, the model estimates the

threat posed as being dictated by the largest of the constituent factors (e.g., an extremely high motivation would in itself be sufficient to assess someone as a threat).

Our quasi-Bayesian method uses a simplified way of reflecting report qualities and reliabilities. Also, in sequential updating it allows for some "stickiness" (reluctance to change). We calculate the update as a weighted sum of the current estimate (prior) and the estimate obtained by a Bayesian update. We refer to this as quasi-Bayesian because the method has the familiar concepts of prior, likelihood function, and posterior, but also has the stickiness effect and other approximations. Also, the prior assessment may be based on little more than vague suspicions, and the likelihood functions are also uncertain. The prototype allows the analyst to compare results with several generic likelihood functions and also allows for contextually specific subjective estimates.

We developed the maximum entropy/minimum penalty (MEMP) approach to supplement the other methods, in part by providing a "conservative" calculation that does not speculate beyond what is known or explicitly assumed. Maximum entropy methods are well understood. For our context, however, much of the relevant knowledge is uncertain, fragmentary, and of questionable reliability, making standard maximum entropy methods difficult. We therefore drew on methods from other fields, notably regularization methods from machine learning, to represent the noncertain information with soft constraints and slack variables. We then generate the probability distributions by minimizing a weighted sum of entropy terms and penalty terms and by then estimating the appropriate weighting factor. The result is to approximate the intent of a maximum entropy method while dealing with the complex, uncertain, and heterogeneous information. The MEMP method considers information from all reports at the same time, rather than using a sequence of step-by-step updates. It deals with flatly contradictory information.

Creating Synthetic Data

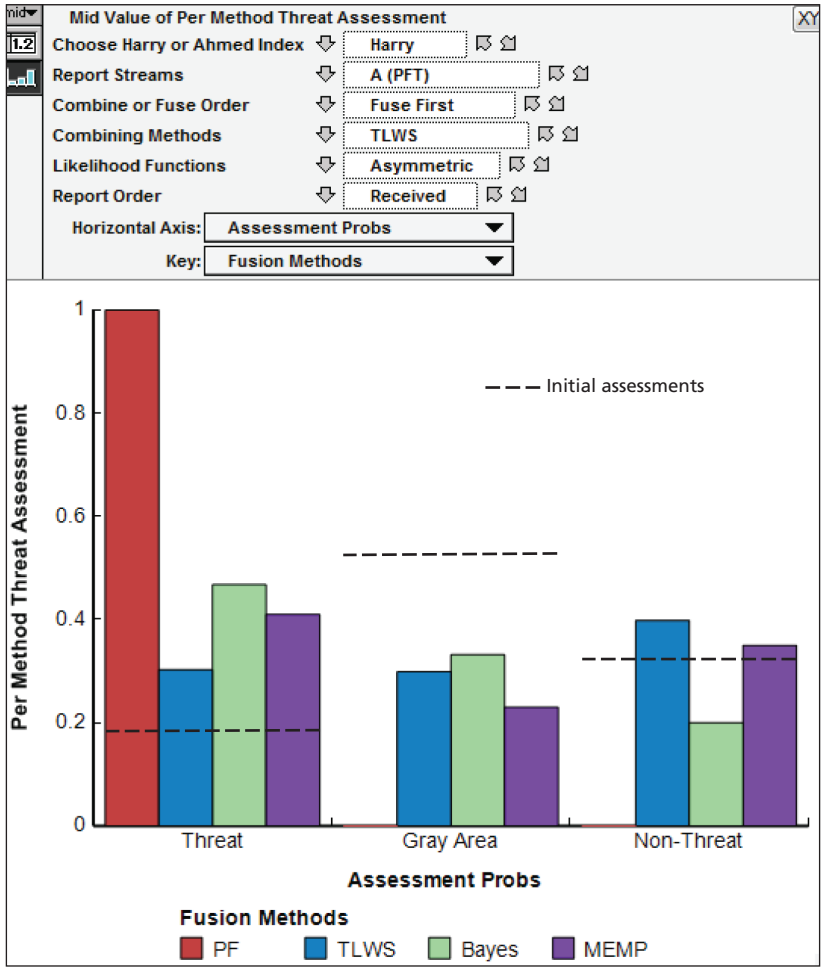
For proof-of-principle experiments, we developed concrete synthetic test cases (about 40 pages of text), expressed in plain but colorful language. These served as "vignettes" in the continuing cases of hypothetical individuals: Harry (a potential right-wing terrorist) and Ahmed (a potential jihadi terrorist). These test cases were designed to stress the prototype by including particular challenges of information representation that we had identified earlier in the project. Thus, they included contradictory reports, reports that addressed only some of the issues necessary to assess threat, and reports that might be deliberately deceptive. The reports came with estimates of credibility, but those estimates were also uncertain. The reports were to be mapped into analytically usable constructs with accompanying expressions of uncertainty. For example, a particular report might say that it was unlikely that Harry was more than modestly motivated for a cause, that he had high or very high capability for terrorist actions, but that he had little spine for anything risky.

The concreteness of the vignettes stimulated finding practical ways to deal with challenges that appeared more formidable in the abstract. Further, it improved communication and mutual understanding of the concepts at issue. Such work with synthetic data is not truly empirical, but it is nonetheless valuable for research. Our use of vignettes is analogous to using stressful scenarios in strategic planning exercises and wargaming.

Initial Experimentation

Late in the research, the prototype platform and the various fusion methods came together. This was gratifying because the complexities had been a worry. Although only limited experimentation was possible in the project, it confirmed feasibility and allowed us to learn a good deal about the methods and their various strengths and weaknesses. Figure S.3 shows illustrative results based on the synthetic data. Each bar shows the assessed probability, after fusion, that the individ-

Figure S.3
Fusion Raises Estimate Suspect Is a Threat



RAND *RR1200-S.3*

ual in question should be classified as a threat, in the gray area, or as a nonthreat. These categories represent an aggregation corresponding to threat levels T (measured on a 0 to 10 scale) of 6 to 10, 4 to 6, and 0 to 4, respectively. Each clump of bars shows how results change as a function of the fusion method indicated by bar colors. The methods

are primary factors (PF), thresholded linear weighted sums (TLWS), Bayesian (called quasi-Bayesian in the text), and maximum entropy/minimum penalty (MEMP). The dashed horizontal lines indicate the corresponding initial assessments (i.e., the "priors"). The initial assessment (directly subjective in our examples) puts the likelihood of Harry being a threat as 0.18, i.e., less than 1 in 5. The fused estimates drawing on additional information change this assessment.

At the top of Figure S.3, we see a set of interface elements called "slicer bars." These show the values of selected other variables. The analyst can explore the effects of changing these values, seeing results across all the possible cases generated by the uncertainties in these variables and choice of fusion method. This illustrates how multivariable uncertainty analysis can be accomplished routinely, rather than being constantly deferred. Reading assumptions from the top down, Figure S.3 shows results for Chapter Three's Harry vignettes. It shows results from only Stream A of analysis. Results are for the fuse-first approach, in which model factors are estimated by fusing across reports, after which threat is estimated by combining the factors with the TLWS method. For the results based on quasi-Bayesian methods, the asymmetric likelihood function is used. The reports are processed in the order received. With these assumptions and the MEMP fusion method, Harry is given a 40 percent likelihood of being in the threat category, about twice the initial estimate (fourth bar in left clump of bars).

The analyst can explore the consequences of different choices for any or all of these assumptions. The example, then, illustrates exploratory analysis. Many uncertainties can be handled, especially with machine methods to help. There is no excuse anymore for not dealing seriously with uncertainties.

For this particular set of assumptions, the probability that Harry is a threat varies markedly, from about 0.3 to 1, depending on the fusion method used. The PF method concludes that Harry is unequivocally a threat, whereas the other methods see Harry's being a threat as much more uncertain. The TLWS method is least alarmist, because it sees no threat unless *all* the factors of the model being used are consistent with the individual being a threat.

The last block of three bars shows related differences for the estimated probability that Harry is a nonthreat. Even if we put aside the PF result, which gives zero credence to Harry being a nonthreat, results vary by a factor of two depending on method.

Ordinarily, analysts hope that calculations done in different ways will lead to the same results: The intent is just to check for errors. In contrast, we are comparing across different methods specifically because we know that they may give different results and because we do not know a priori which method is most nearly "right." When the results do disagree, as in this case, it is necessary to look deeply at the particulars of the case to better assess which methods and assumptions make the most sense. We also need to understand how sensitively results depend on particular data, especially if uncertain or even suspicious.

Our experiments demonstrated the flexibility of our prototype system. In particular, we illustrated the ability to vary

1. fusion method
2. causal model
3. relative order of fusing model-factor information and fusing threat estimates
4. subjective assessments of credibility and reliability
5. likelihood function
6. heuristic cut-off rules on quality
7. thresholds
8. whether and when to fuse across streams of analysis.

The experiments also demonstrated that the prototype *could* be comprehensible to analysts with suitable background, despite its complexity. That said, the experiments also revealed shortcomings, such as where further simplifications or additional explanation features were needed.

Conclusions and Next Steps

Our study demonstrated the ability to construct a suite of fusion methods and apply them usefully to some relatively complex synthetic data, accounting for uncertainties in data and about which methods to use. As expected from the design, in some cases, fusion elevated the estimated likelihood that an individual was a threat, suggesting that fusion can help improve threat detection. In other cases, it elevated the probability that an individual was *not* a threat, suggesting that fusion can help reduce false alarms. In still other cases, the uncertainty analysis demonstrated ways to view the information that would be exonerating if confirmed.

Our research to date has been basic in nature and should be seen as first steps in a process. We were encouraged by results but cannot as yet assess what could be accomplished with real data, nor compare value with the current baseline of essentially subjective analysis. We hope in future work to be able to do the following:

1. Experiment more with the prototype to better assimilate its lessons.
2. Sharpen and extend our methods (e.g., to include a Bayesian-network method).
3. Fill in theoretical gaps (e.g., understanding how best to use empirical base-rate data and the results of data-driven fusion exploiting data mining and machine learning).
4. Develop a more mature prototype platform that can be shared.
5. Use more comprehensive and representative data sets informed by real-world and exercise experience in the government.
6. Conduct experiments with analysts experienced in current-day high-level fusion operations in government.

If such work is successful, then it will be time for relevant agencies to move toward applied development. The value of the approach will then depend also on refining the quality of the models and methods and educating analysts in their use. Even then, of course, effectiveness in real-world applications will depend on the quality of the information available.

Acknowledgments

A particular pleasure in this research was going back to original works by pioneers such as Edwin Jaynes. We benefited from a number of texts and current literature on decision analysis, inference, Bayesian nets, machine learning, and information fusion in other domains. The report is much stronger because of reviews by Kathryn Blackmond Laskey (George Mason University) and RAND colleagues Joel Predd and Andrew Liepman.

CHAPTER ONE

Introduction

Military and other government organizations have put substantial effort into detecting and thwarting attacks such as those by suicide bombers or involving improvised explosive devices (IEDs). Such attacks may be against military or government installations in the United States or abroad, against civilian infrastructure, or any of many other targets. An element of thwarting such attacks is observing suspicious individuals over time with such diverse means as scanners and other devices, travel records, behavioral observations, and intelligence sources. Such observations provide data that are often both complex and "soft"—i.e., qualitative, subjective, fuzzy, or ambiguous—contradictory, and even deceptive. The problem, then, is how to combine the data to form a realistic assessment. We refer to this as *heterogeneous information fusion*. In this report, we summarize research on methods for heterogeneous fusion. The context is counterterrorism, for both military and civilian applications, but the ideas are also applicable in intelligence and law enforcement.

This report describes basic research on such fusion, with potentially broad contributions to counterterrorism, intelligence and law enforcement.[*] All fusion is heterogeneous to some degree, but we have

[*] Information fusion is "the study of efficient methods for automatically or semi-automatically transforming information from different sources and different points in time into a representation that provides effective support for human or automated decision making" (Boström et al., 2007). For fusion methods in physics-centric problems, see Klein (2004). For a review of multisensory data fusion, see Khaleghi et al. (2013). The primary journals are *Information Fusion* and *Journal of Advances in Information Fusion*. Chapter Four points to prior research

in mind multimethod fusing of complex information from diverse sources. Significantly, we also have in mind fusion processes that routinely characterize uncertainty and—more unusually—report how fusion results depend on assumptions and analytic methods so that iterative analysis can be more discriminate and convergent.

Objective

The context of our research is assessing the threat of terrorism posed by individuals or groups under scrutiny. Broadly, the ultimate objectives, which go well beyond what we sought to accomplish in our study, include improving the detection of terrorists, reducing false alarms, and exonerating those who have been inappropriately suspected. The need for such improvements was discussed in an earlier study (Davis, Perry, et al., 2013). The stakes are high for both the defense of the nation and the protection of civil liberties and commerce (National Research Council, 2008). Based on public records, we can say in retrospect that information existed that *might* have allowed detection and interdiction of the attacks of September 11, 2001; the apparent sender of the 2001 anthrax letters went undetected for a decade, while another individual was pursued relentlessly; and the number of people on watch lists is reportedly on the order of 1 million.*

Our particular research had the narrower objective of designing and implementing a research prototype system for uncertainty-sensitive heterogeneous information fusion with the potential for assisting progress toward these broad objectives. Success with our prototype would be only one step, but it would justify next steps toward applications.

---

relevant to softer aspects of heterogeneous fusion (e.g., aggregation of expert forecasts and multivariable Bayesian nets) where data are sparse and uncertain.

* For the 9/11 case, see National Commission on Terrorist Attacks (2004) or Bergen (2013). The anthrax case remains controversial (Shane, 2015). The watch-list estimate is from the American Civil Liberties Union (2015) (as of November 19, 2015: https://www.aclu.org/terror-watch-list-counter-million-plus).

Research Strategy and an Idealized Fusion Process

What prototype system should we design and build? After reviewing the issues and drawing on the scientific literature, we settled on a strategy of building a system with a number of particular attributes. The system in question should do the following:

1. Address matters probabilistically rather than seeking point conclusions.
2. Take a mixed-method approach to information fusion.
3. Allow for parallel, competitive, and iterative streams of analysis.
4. Employ causal models rather than just statistical data.
5. Routinely use exploratory-analysis methods to deal with uncertainty affecting models, process, and assumptions.

Such a system should contribute to achieving the broad objectives mentioned above. Fusion, especially with results expressed probabilistically to retain more information, might be able to identify threats that were not recognized before the fusion and, conversely, to identify individuals as nonthreats despite the existence of some apparently adverse information. There was reason to expect that fusing across methods and having competitive analyses would be helpful. Using causal models (rather than statistical models describing correlations) should improve substantive reasoning, inference from fragmentary information, and explanation. Given the many uncertainties and choices to be made in information fusion, systematic exploratory analysis should help avoid premature or overly confident best guesses and should allow finding conclusions that stand up well across many of the uncertainties.

To be sure, the *real-world* value of such a system, if feasible, would depend on all the details: the quality of the methods, models, data, and analysts. Our practical objective was the first step: to achieve technical proof of principle using first-cut methods, illustrative models, and appropriately designed synthetic data. Could we design and pull all these pieces together into a comprehensible and manageable system for analysis? The strategy implied numerous tasks, notably (1) representing complex information, (2) constructing illustrative causal models,

(3) constructing a diverse set of fusion methods, (4) constructing synthetic data that would challenge the prototype, and (5) integrating the elements in a research-level prototype computer platform.

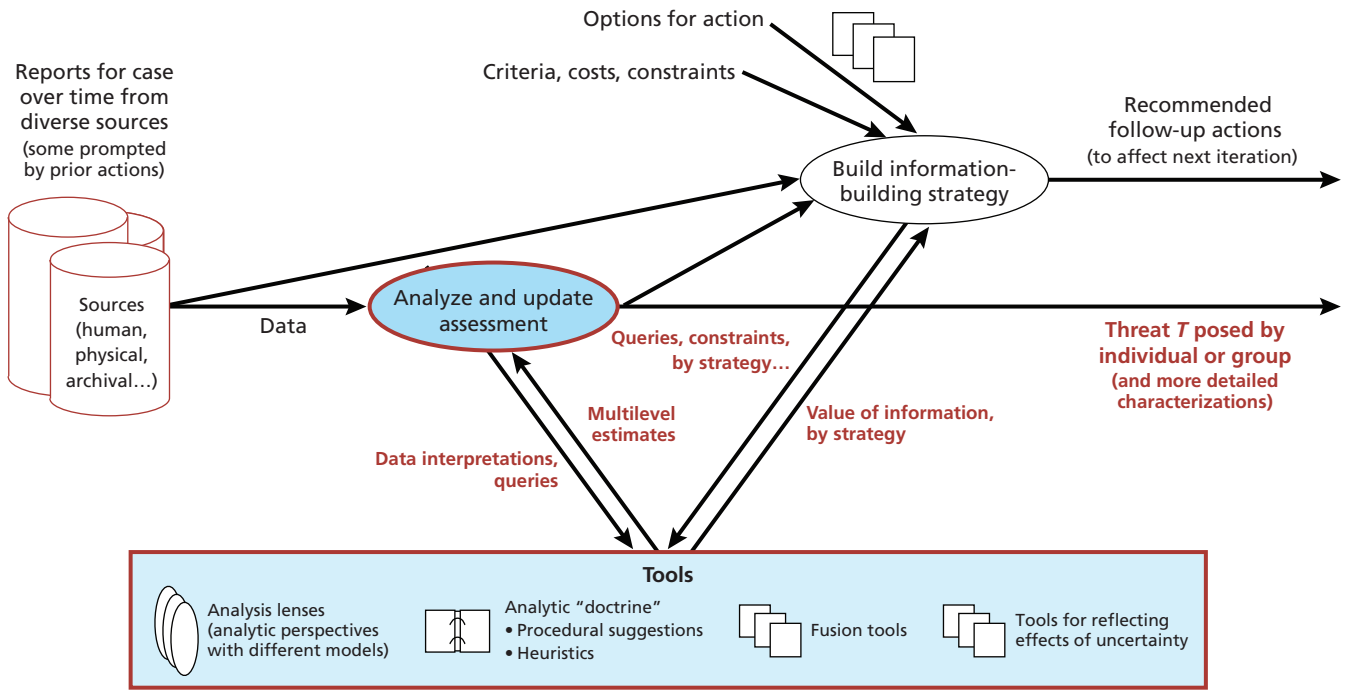
Figure 1.1 depicts our vision of an ideal process for uncertainty-sensitive heterogeneous fusion. Many sources of data (left) provide inputs for assessments by exploiting a variety of tools (bottom shaded area), which involve different analytic lenses (i.e., different analytic perspectives, each with its own model), fusion methods, and methods for uncertainty analysis. As at the top right, the process informs information-gathering: What information, if obtained, could narrow uncertainty in useful ways? What information-gathering is feasible given legal and other criteria and costs? The output (right side) is an assessment of the threat T posed by the individual or group under scrutiny, other characterizations (such as motivation and capability for terrorist acts), *and* recommended follow-up actions. The initial research reported here focuses strictly on the items highlighted in red, analyzing and updating assessments, but the larger idealization is important to keep in mind.

Analytic Architecture for Uncertainty-Sensitive Fusion

How could we move toward the vision of Figure 1.1? We designed a generic analytic framework for uncertainty-sensitive heterogeneous information fusion, shown in Figure 1.2. In this, evaluating the threat posed by an individual or group involves multiple reports and multiple streams of effort by analyst teams with their own models, methods, and judgments. All this may occur within or across responsible organizations. At any given time, fusion might occur across the reports and analysis streams to generate an integrated threat assessment. Subsequent paragraphs provide initial background on the concept of "threat," what we mean by models, and how we deal with uncertainty.

Threat is represented here as a variable T between 0 to 10, with higher numbers being more worrisome, as indicated in Table 1.1. Often we use a corresponding discrete scale with values 1, 3, 5, 7, and 9. The ranges 0–2, 2–4, 4–6, 6–8, and 8–10 are referred to qualitatively

Figure 1.1
Idealized Fusion



RAND *RR1200-1.1*

Figure 1.2
Initial View of Heterogeneous Fusion Process

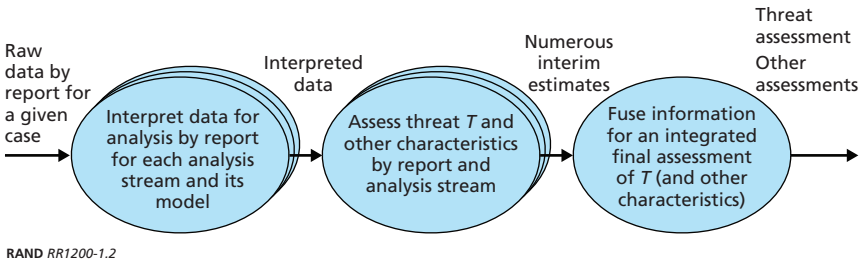


Table 1.1
Threat Levels

Level or Tier	Nominal Score (Range)	Qualitative	Example
1	1 (0–2)	**Very low**: not a threat or otherwise not worth follow-up action	Someone exonerated (score 0), no more of a threat than average person
2	3 (2–4)	**Low**: probably not a threat but merits modest follow-up action	Someone associated with a low-consequence action or one of highly doubtful feasibility
3	5 (4–6)	**Medium**: merits follow-up actions to improve information	Someone associated with a possible medium-consequence action
4	7 (6–8)	**High**: merits follow-up actions with expectation of likely surveillance, arrest, or interdiction	Someone associated with a possible high- or medium-consequence action
5	9 (8–10)	**Very High**: merits maximum response, with regular updates at top level and extensive alerts down the line	Someone associated with a possible high-consequence action and a non-trivial possibility of success

as very low, low, medium, high, and very high, respectively. Chapter Two provides more details. As discussed in Appendix A, the assessed threat level depends on the kind of attack with which an individual might be associated and the plausibility of his attempting and to some extent succeeding. The assessment is specific to the relevant agency: For

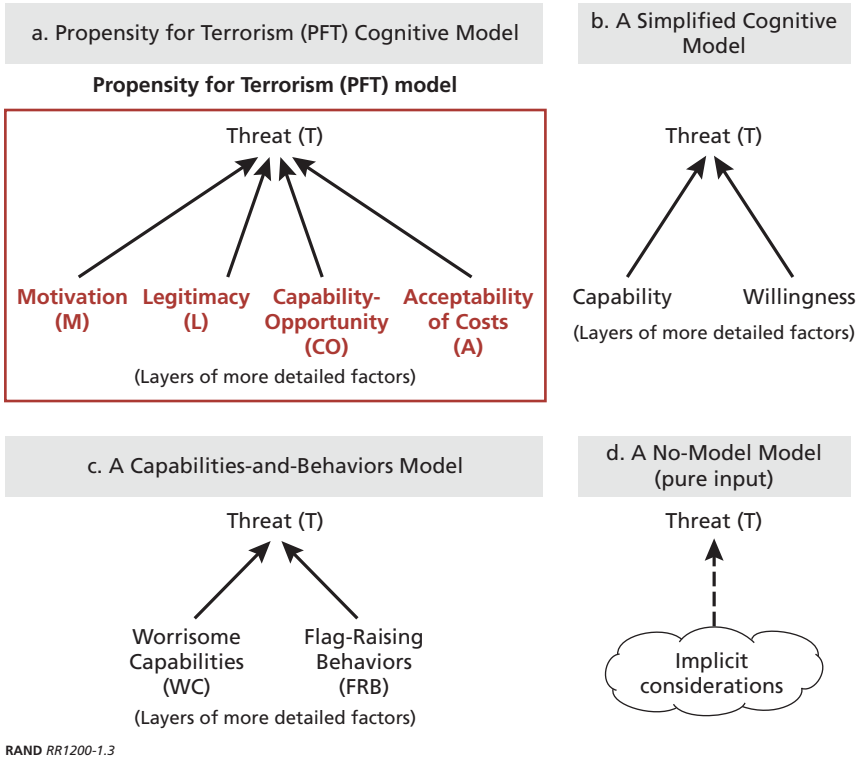
example, someone intent on armed robbery is seen as a higher threat by police than by a national counterterrorism agency.

Models play a crucial role. We emphasize *causal* models (i.e., those expressing cause-effect relationships), rather than statistical correlations. Figure 1.3 illustrates schematically what we mean. Pane (a) uses the Propensity for Terrorism (PFT) model described in Appendix B. Although by no means definitive, it has a base in social-science research and sees T as a function of four variables or factors: motivation (M) for a cause or activity, legitimacy (L) perceived in committing acts of terrorism, capability-opportunity (CO) for the terrorist acts (having the skills required and access to the means needed for the attack), and acceptance (A) of costs (such as the risks of death or capture). Three of the variables relate to cognition and reasoning. The model of pane (b) considers only willingness and ability. That of pane (c) considers only worrisome capabilities and flag-raising behaviors, with no attempt to reflect psychology. Pane (d) indicates a "no-model model"—i.e., one in which judgments about threat posed by the individual are made intuitively. These are only four of many possibilities.

Uncertainties must be addressed all along the threat-assessment process, often applying subjective judgments. The uncertainties can be bewildering or paralyzing, but great strides have been made in dealing with uncertainties effectively (Morgan and Henrion, 1992; Laskey, 1996, Davis, 2003, 2012; Lempert et al, 2006). Doing so was one of our primary objectives. As part of this, we draw sharp distinctions among classes of uncertainty as summarized in Table 1.2. First, we address both model uncertainty and input uncertainty. Analysis may be inaccurate because of a poor model, poor data for a model, or both.

No comprehensive and reliable methods exist for dealing with model uncertainty, but much can sometimes be accomplished by considering a well-chosen variety of model structures, as in using alternative cognitive models of the adversary or having competing teams use their preferred models. It is easier to deal effectively with input uncertainties (rather than model uncertainties), whether by varying deterministic parameter values, representing variables probabilistically, or both.

Figure 1.3
Schematics of Alternative Causal Models



From the outset, we regard the threat characterization T as probabilistic, rather than considering just the so-called best estimate. We use the term *probabilistic* rather than *random* because the uncertainties in question *typically* reflect shortfalls in our knowledge, rather than the existence of random processes. We do, however, allow for random processes, such as whether a potential terrorist will happen to find a skilled terrorist organization to join.

Another issue related to uncertainty has to do with "stories." We want to be explicit where subjective considerations play a major role in information fusion. This requires recognizing what are variously referred to as stories, narratives, or mental models that humans use to

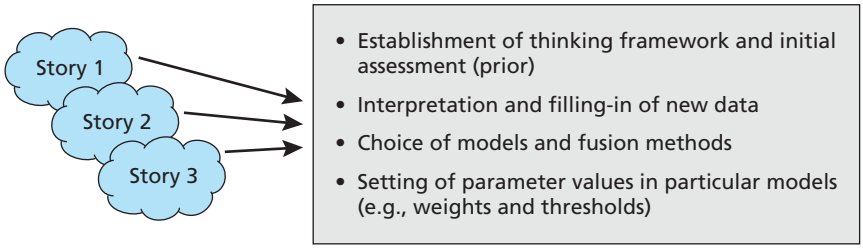
Table 1.2
Ways to Reflect Uncertainty

Type of Uncertainties	Methods
Model uncertainty (structural uncertainty)	
	Use alternative models
	Use parameters to change effective model structure
	Use probability distributions to represent random (stochastic) processes
Input uncertainty	
Parametric	Vary discrete values of deterministic parameters
	Represent knowledge gaps with probability distributions
	Use both parametric variation and probability distributions
Inherent (random effects)	Use probability distributions to represent random variables

make sense of data and connect dots.* Some stories may be sanguine, while others may impute complex causality as in conspiracy theory. Sometimes, belief in a story is bad—as when the story incorporates inaccurate beliefs due to prejudice, inexperience, or failure to understand the normal frequency with which observed events occur. Other times, it is good—as when an expert drawing on intuition quickly sees patterns that turn out to be accurate. Figure 1.4 indicates schematically that the stories at play in an analysis have numerous effects. They affect how we interpret data and extrapolate. Analysis should therefore make such considerations explicit for recognition, review, debate, and either resolution or hedging. The methods for doing so are the same as those discussed above for dealing with uncertainty. When stories are particularly important in fusing information, the analytic engine of Figure 1.2 may best be seen as driven by higher-level questions or beliefs.

* This is related to "explanation-based decisionmaking" (Pennington and Hastie, 1988, reprinted in Goldstein and Hogarth, 1997) and the need for what the CIA calls the evaluation of alternative hypotheses (Heuer, 1999). Stories can also play a major role in the reasoning of terrorists themselves, not just the investigators (Atran, 2010).

Figure 1.4
How Stories Influence Analysis



RAND *RR1200-1.4*

About This Report

The remainder of the report is structured as follows. Chapter Two describes concepts, methods, and a research platform for experimentation. Chapter Three describes test cases and related vignettes. Chapters Four and Five discuss different fusion methods using examples from Chapter Three. Chapter Six shows some initial results and comparisons across methods. Chapter Seven gives conclusions and recommendations. We include three appendixes: Appendix A defines what is meant by a terrorist threat, Appendix B describes the factor-tree Propensity for Terrorism model used throughout this report, and Appendix C provides details for some subtleties of Bayes calculations. We have more documentary detail in unpublished materials.

CHAPTER TWO

Concepts, Methods, and a Research Platform

This chapter begins by discussing the prototype research platform, development of which was an important part of our research. Most of the chapter is an introduction to the concepts and methods used.

Research Platform

We needed a prototype research platform to help us develop and understand concepts and methods. Concreteness helps in understanding subtleties and in stimulating solutions to difficult problems. We largely implemented the framework model of Chapter One in Lumina Descision Systems' Analytica software, which is based on a high-level functional language well suited to expressing array mathematics and the related issue of dealing with uncertainty. It features visual modeling with influence diagrams and has extensive built-in features for probability, statistics, and uncertainty analysis.* Our intent was that even the prototype platform should be largely understandable by technically educated analysts who are not expert programmers and that the linkages between the computer program and the underlying math-

* Analytica is a product of Lumina Decision Systems. Its influence diagrams need not be probabilistic, as in Bayesian-net and influence-net work using, e.g., the Netica program of Norsys Inc. They also differ from the causal-loop diagrams of System Dynamics (Sterman, 2000). See a text on uncertainty analysis (Morgan and Henrion, 1992) and the Lumina website (as of November 19, 2015: www.lumina.com) for examples. As an exception to our use of Analytica, we implemented entropy maximization, as described in Chapter Five, in Microsoft Excel, although with the intention of later integration.

ematics should be clearer than with many programming languages. This was not fully achieved with the prototype, because implementing the flexible architecture required advanced programming methods that undercut program clarity. We plan a more mature version that will be more transparent and suitable for direct re-use and that will serve as a rigorous specification for re-programming into other computing environments if desirable.

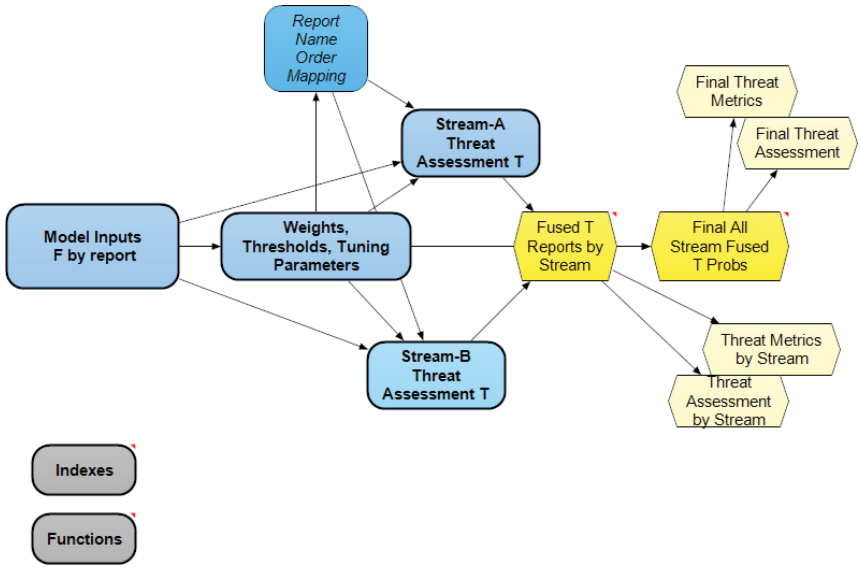
Figure 2.1 is a screenshot of the top-level modules and how they influence each other. The left-most module contains input information about the individuals being assessed. The next module contains weights, thresholds, and tuning parameters used by the fusion algorithms. The fusion analysis occurs separately in the Stream-A and Stream-B modules. A vast number of outputs are possible, but the rightmost objects (the yellow ones) are the most important. The gray modules contain index variables and mathematical functions used across all of the platform's nodes, thereby improving comprehensibility and complexity management.

Representing Qualitative and Quantitative Information Consistently

One of our challenges was to represent information that may be qualitative or quantitative, crisply or vaguely expressed, complete or fragmentary, and certain or uncertain—including contradictory and possibly false information. We first discuss simple examples, moving to more complicated ones later.

Much of the relevant information is qualitative, as when an individual is said to be highly motivated. We define qualitative concepts on a common 0-to-10 scale with equal-spaced values that enables mathematical operations that would not be valid for ordinal scales. We then use numbers and qualitative expressions interchangeably when referring to values, although using the numerical versions for calculations. Table 2.1 shows the five-level discrete version of this. The second and third columns explain the sometimes-confusing distinction between the level or tier number and typical value in a tier.

Figure 2.1
Initial Version of the Research Platform



RAND *RR1200-2.1*

Table 2.1
Representing Qualitative Variables

Qualitative Expression	Tier	Nominal Value	Continuous Range
Very Low	1	1	0–2
Low	2	3	2–4
Medium	3	5	4–6
High	4	7	6–8
Very High	5	9	8–10

NOTE: More precisely, using set notation the tier-ranges are [0,2), [2,4), [4,6), [6,8), and [8,10].

Expressing Simple Uncertainty

The significance of uncertainty in variable X may be examined by changing its value in deterministic calculations. For example, we might let X have values in the set {1,3,5,7,9} and conduct the analysis for each. Or we might just use the bounding cases of X=1 and X=9.

An alternative is to represent the uncertainty by defining a probability distribution Pr(X). This is not the same as treating X as a "random process" in the usual sense of that term. In most cases, the relevant uncertainty has to do with shortcomings in knowledge, rather than some random processes in nature. An individual may actually pose a definite level of threat. However, that level is uncertain to us. The probability distribution reflects that uncertainty.

> Subjective probabilities reflect uncertainties of knowledge rather than random processes.

Hybrid methods for dealing with uncertainty use probabilistic methods for many uncertainties but use deterministic parametric methods to highlight the implications of others. Such methods have been used extensively (Davis, 2003, 2012, 2014a).

Although allowing use of arbitrary forms, including the familiar normal, log normal, and beta distributions, we standardized on combinations of triangular and uniform distributions for continuous variables and on the simple five-value discrete distribution of Table 2.1. Doing so meant forgoing some analytic elegance (e.g., closed-form solutions are sometimes possible with normal or beta distributions), but it was necessary for generality. We also developed protocols and model functions for moving between continuous and discrete representations easily.

Our standard choices also lend themselves well to eliciting subjective information. In developing protocols for interpreting such subjective inputs, we drew on our personal experience, the literature (e.g., O'Hagan and Oakley, 2004), and intelligence-community guidelines based on psychological research (Heuer and Pherson, 2014; Heuer, 1999).

Expressing More Complex Forms of Uncertainty

We needed fusion methods to handle complex expressions of information (Davis, Perry, et al., 2013; Perry, 2011). To better understand the range of challenges, we drew on the theoretical literature for Bayes, Dempster-Shafer, and Dezert-Smarandache methods (Shafer and Dempster, undated; Shafer, 1976; Shafer and Pearl, 1990; Smarandache and Dezert, 2009a, 2009b). We were also influenced by Judea Pearl's work on causality (Pearl, 2009). A challenge was whether we could represent all the expressive forms discussed by these authors, including disagreements and contradictions. The next sections address these issues.

Complex Disjunctions, Conjunctions, and Complements

Complex information may consist of all combinations of disjunctions, conjunctions, and complements—i.e., or's, and's, and not's. That is, reports in our context may include logical constructs such as "he's up to this *or* that, but nothing else"; "I don't know what he's up to, but it's definitely *not* that"; and "I really think he's up to this *and* that only." The italicized words illustrate disjunction, conjunction, and complementarity (i.e., "or," "and," and "not" relationships). Statements may also be made at different levels of detail, referring to different factors in a model or to different value levels of a given factor.

To illustrate the issues systematically, we constructed test cases, such as the following, in informal language (the full test cases were more elaborate, as discussed in Chapter Four):

1. "I'm not sure what he's up to, but he definitely has no access to bomb-making ingredients, so he can't be making a bomb."
2. "This newly formed group is either highly motivated to conduct some terrorist attack or it's just a naive disgruntled bunch."
3. "He's definitely motivated to conduct some terrorist attack, and he has access to the means to carry it out."
4. "This group is clearly not motivated and has no means to carry out an attack."
5. "This guy is no threat at all—his talk is sheer bravado."

Using the PFT model for concreteness, we interpreted the five cases as statements about the threat factors (motivation, legitimacy, etc.) and their strength levels (very low to very high or, numerically, 1, 3, 5, 7, 9). We then expressed the statements mathematically using the logical connectors mentioned above. We found that we were able to represent the complex information with the mechanisms adopted: causal threat models, multifactor subjective probability distributions, multilevel hypothesis sets for each model factor, and the logical connectors.

Contradictory Evidence

Contradictory evidence is a special problem in information fusion. It can arise within a single report, as when a source indicates that an individual "is either highly motivated or putting on an act" or when reports contradict each other. Perhaps one report claims that the subject sincerely believes in the legitimacy of the cause based on his postings on a jihadist blog. The second claims that the individual has naively agreed to write elaborate justifications for terrorism as a favor to his good friend, but is well aware of how bankrupt the cause is and would never participate in an act of terrorism.

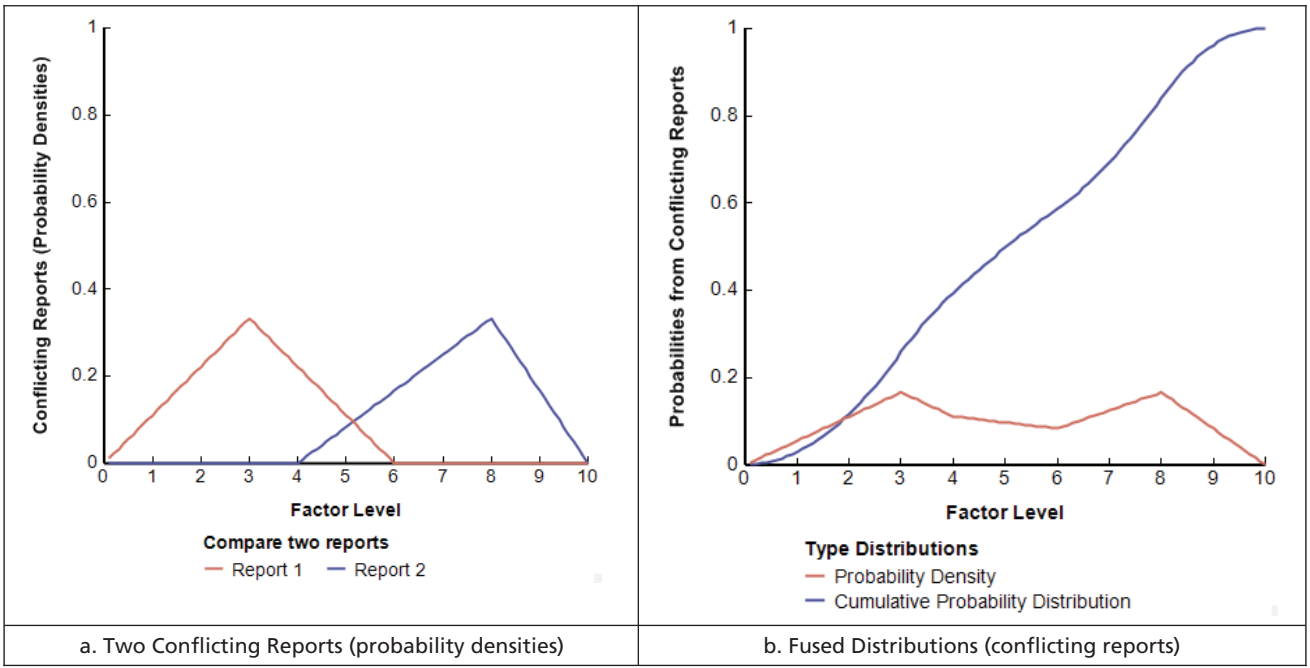
Conflicting reports can be handled in the same way regardless of source, fusing them with algebraic methods discussed in Chapter Four. The fused distribution function is not the "average" distribution, but rather something that preserves the structure of the competing reports, as in Figure 2.2. Fig. 2.2(a) shows the probability densities for the conflicting reports. Fig 2.2(b) shows the fused probability density and cumulative probability distribution. Note that the probability density (red curve of pane b) is bimodal.*

Table 2.2 summarizes how we dealt with a variety of representational challenges.†

* We process such input distributions before the fusion operations to assure that they give at least some probability mass to all portions of the [0,10] range. This avoids certain errors in Bayesian fusion.

† See Khaleghi et al. (2013) for a taxonomy of information-fusion issues from a more usual perspective.

Figure 2.2
Representing Uncertainty for Complex and Conflicting Inputs



a. Two Conflicting Reports (probability densities)

b. Fused Distributions (conflicting reports)

RAND *RR1200-2.2*

Table 2.2
Challenges and Responses in Representing Information

Challenges	Example	Representations
	Dealing with Uncertainties	
Uncertain knowledge versus randomness	Information on a factor may be uncertain because of limitations in knowledge or, sometimes, because attribute is changing, perhaps randomly.	Probability distributions for characterizing uncertain deterministic knowledge and for reflecting random processes.
Indirect and fragmentary knowledge	Is he a threat? One report: "He seems motivated"; a second report: "He has capability." Neither report covers all the factors affecting threat.	Threat assessment models using separately observable factors and filling in missing information with, e.g., empirical base rates.
Correlations	Information that is seemingly independent may be highly correlated.	Models with independent observable factors. Mechanisms for representing correlations. Sensitivity testing.
	Combining Information	
Disjunction (or)	Report: "Motivation is very high *or* very low, but not both."	Bimodal distributions. Fusion methods that preserve disjunctions.
Conjunction (and)	Report: "Motivation is high and so is capability."	Multi-attribute (multi-factor) distributions.
Complement (not)	Report: "He is not a threat; he has no motivation."	Assign probabilities accordingly.
Inequality	Report: "Motivation is more than nothing but less than very high."	Algebraic inequalities connecting distributions, e.g., very low < motivation < very high.
Likelihoods (conditional probabilities)	Given previous information, how should new information be interpreted?	Approximate Bayesian likelihood functions, where applicable, but with uncertainty analysis using alternative functions.

Table 2.2—Continued

Challenges	Example	Representations
	Dealing with Conflicts, Disagreements, and Deception	
Disagreement to be resolved	First source says sees definite motivation; a second report sees the absence of motivation.	Fusion methods that preserve disagreement. Quality assessment of reports.
Disagreement to be reflected or reduced	Sources give estimates of motivation that vary from very low to very high, with no pattern.	Fusion methods that average or that estimate underlying convergence.
Deception about knowledge	Sources disagree on motivation but some are fabricating their stories.	Distinct reports assessed subjectively for quality and reliability.
Deception with intent to cause mis-assessment	Sources disagree; some are deliberately misleading: personal enemies may plant false negative information and friends may plant false positive information.	Distinct reports with associated reliabilities. Fusion methods that preserve conflicts. Alternative stories (higher-level "explanations").

Dealing with Probabilistic Correlations

A fundamental challenge was finding ways to deal with the probabilistic dependencies among variables potentially affecting the threat estimate. As background, logically independent variables may be probabilistically correlated. Consider a function F of variables $\{X_i\}$. The X_i are *logically* independent—if it is possible to change each of them without effect on the others. The variable X_i has an independent effect on F if, for at least some portion of the domain of X values,

$$\left(\frac{\partial F}{\partial X_i}\right) \neq 0 . \tag{2.1}$$

Meeting this condition does not mean that the *values* of the X_i are equally important or that the X_i are probabilistically independent. Further, even if the model represented by F is deterministic (producing the same results each time it is evaluated for a given set of inputs), the input values may be uncertain. Uncertainty analysis can vary those inputs parametrically, represent them probabilistically, or a combination. The parametric approach is often convenient and has been a mainstay of

RAND's *exploratory analysis* (Davis, 2003, 2012, 2014a) and *robust decision making (RDM).*[*] After running the parametrics, however, one must then ponder about how seriously to take good and bad cases. In doing so, one must worry about probabilistic correlations, because the various cases generated by the different combinations of variable values are not equally probable.

Using probability distributions instead of parametric variation introduces some mathematical difficulties. The joint probability distribution $\Pr(X_1, X_2, \ldots, X_N)$ is, in the general case, formidable. The challenge is how to deal with these correlation-related matters in a way that is reasonable but mathematically tractable. This requires domain knowledge.

Our strategy was a defense in depth:

1. *Define variables* (factors) of our model so that they are probabilistically independent to the extent possible.
2. *Use diverse fusion methods*, including some that are relatively insensitive to correlations (such as the primary factors method described in Chapter Five).
3. *Permit exceptions* by allowing the analyst to directly specify correlated relationships when there is reason to believe that raw inputs have such correlations.
4. *Hedge* with uncertainty analysis measuring the *potential* effects of correlations by parameterizing correlation (Davis and O'Mahony, 2013).

We followed this strategy when using the prototype models discussed earlier (Figures 1.3, panels a and b) and in Appendix B.

[*] See the RAND Corporation's web page on the topic; as of November 19, 2015: http://www.rand.org/topics/robust-decision-making.html

Subjective Estimates of Quality, Credibility, Salience, and Reliability

Throughout our fusion calculations, it is necessary to apply weights or otherwise distinguish among factors, reports, or streams. In some cases, weights can come from empirical social-science research, but more often they must be subjective entries. The quality ascribed to a given report should depend on the credibility of the source and how salient the report's evidence is for what is being estimated. Some reports of equal apparent quality when judged in standardized ways will be contradictory. It is then necessary for the analyst to specify a tie-breaking multiplier. All of these inputs are between 0 and 1. The net result of quality times the multiplier is what we refer to as *reliability*. This is translated into a normalized weight internal to the calculations. Table 2.3 illustrates the effective data structure for a given individual if there are two streams of analysis, two reports in each, and four model factors. The inputs are to be entered in the yellow field. The analyst may input quality directly or specify it as the product of credibility and salience if those are inputted. The analyst producing a given report should provide the quality-related assessments. However, a subsequent

Table 2.3
A "Flat" Data Table for Report Quality and Reliability Inputs

Item	Stream	Report	Factor	Credibility (optional)	Salience (optional)	Quality	Subjective Multiplier	Final Weight
1	A	1	F_1					
…								
4	A	1	F_4					
5	A	2	F_1					
…								
8	A	2	F_4					
9	B	1	F_1					
…								

NOTE: The ellipses indicate that some rows are omitted for brevity.

fusion analyst must be able to override them by entering a subjective multiplier.

The reader may be dismayed at the number of subjective inputs, but being explicit about what is often implicit can mean specifying a *lot* of things. This creates a dilemma, because relatively simple models are often preferable to complex models for reasons of transparency, comprehensibility, agility, and organizational acceptance. However, results sometimes depend on details. This is a general dilemma in modeling and analysis. We did not pursue it in this project, but it will be necessary to do so in future work. A key in mitigating the dilemmas is using multiresolution modeling (MRM), which allows for inputting at different levels of detail (Davis, 2012, 2014a). For example, with respect to Table 2.3, the platform could accept single inputs at the level of the last column rather than demanding multiple inputs for each report and factor. Sometimes, that would sacrifice little. Other times, details are important (as when some reports are nearly worthless with respect to some model factors and very good with respect to others).

CHAPTER THREE

Creating Synthetic Data: Vignettes for Testing

Approach

We developed concrete synthetic test cases expressed in plain but colorful language. These took the form of vignettes in the continuing cases of hypothetical individuals: Harry (a potential right-wing terrorist) and Ahmed (a potential jihadi terrorist). We carefully designed the vignettes to pose particular challenges of information representation and fusion (i.e., the vignettes included ambiguous, incomplete, contradictory, fuzzy, deceptive, and subjective information). Thus, while our experimentation was limited to the Harry and Ahmed cases, those were deliberately stressful for methodology testing.

The concreteness of the vignettes improved communication and helped in achieving mutual understanding of the concepts at issue. Also, it stimulated finding practical ways to deal with challenges that, in the abstract, appeared formidable. In many respects, the approach was analogous to detailed scenario development for wargaming and other strategic-planning exercises. However, it was more tightly structured, because the intent was to test methods for rigorous analysis. Thus, the approach has analogies with the practice of finding appropriately comprehensive and stressful test cases in capabilities-based analysis (Davis, 2014a). The result was a methodology that we hope will be a baseline for future work in heterogeneous information fusion.

Our imagined context was a hypothetical National Intelligence Agency (NIA) monitoring activities of seemingly high-risk individuals and groups with information from a wide variety of sources and data types, ranging from digital files about the individuals' history, travel,

and so on, to reports from field agents, local police, closed-circuit camera systems, and call-ins from the public. NIA analysts receive these data, much of them raw, and—when possible—elicit more detailed information by interviewing an agent. The NIA analysts must then interpret the raw information and translate it into an analytically useful form, being scrupulous in representing uncertainties and in recording subjective judgments or choices so that they can be revisited if necessary.

As those familiar with high-quality scenario development for analysis will appreciate, developing such synthetic data required a significant effort. Early versions of the vignettes and interpretations generated about 60–80 pages of information (about 20,000 words). After internal discussion, review, and iteration, we reduced that to a set of short reports totaling perhaps 40 pages and also condensed the data to a few complex tables.

The reports had a structured format driven by analytic requirements and a desire to capture subtleties. In our experiments, all reports focused on behavioral information interpreted as evidence about the four factors of the PFT model described in Appendix B: motivation (*M*) for the cause or activity, perceived legitimacy (*L*) of terrorist violence, capability-opportunity (*CO*), and willingness to accept (*A*) the costs of action. Each report has four sections:

1. Background: why the individual and group have come to the attention of the authorities and, analytically, a prior assessment of the threat posed (with uncertainties).
2. Reports: information, sometimes fragmentary, from agents or other sources. They are narratives describing the raw information and the agents' interpretations. The platform uses only a report's factor-level information or the directly estimated *T*, but not both. The choice of which to use is based on reliability considerations.
3. Quality: credibility and reliability. The analyst must assess the credibility of the agent reporting and the salience of each report, producing an assessment of quality. In addition (either at the time of the report or later, during the fusion process), it may be necessary to subjectively assess each report's *relative* reliabil-

ity. This is necessary because some reports that appear to be of equally high quality (trusted sources, carefully defined information, etc.) may be in conflict. Ultimately, the fusion analysts must decide how much relative emphasis to put on each report.

4. Probability estimates: probability distributions for each model factor expressed as triangular, uniform, or some combination of uniform and triangular distributions, and—in some cases—the source's direct probabilistic characterization of threat.

The two sets of vignettes that we developed for experiments were both set on U.S. soil. They deal, respectively, with potential right-wing threats (Harry Smith and the Slammer organization) or violent religious extremists (Ahmed al-Hiry and the al-Hasqua jihad movement). For brevity, we merely show excerpts in what follows, enough to convey a sense of character.

Harry Smith and the Slammers

Excerpt from initial narrative:

> Harry Smith is an angry, down-on-his-luck day laborer in his mid-30s. His anger is increasingly focused on the "system" that has kept him from succeeding in life. He lives and sometimes works in the city of East Crane, a large city on the East Coast. Unfortunately, East Crane has not benefited from the recent economic recovery enjoyed by other U.S. cities, and unemployment is still high. The out-of-work population in the city is increasingly disaffected, and some people are resorting to crime. More alarming is a recent surge in anti-establishment radical groups that attract this segment of the population, including the notorious Slammers Club composed of anarchists who advocate the overthrow of government. Their tactics include terrorist acts against government facilities. Their recent activities have brought them to the attention of state and national intelligence organizations.

The NIA's Prior Assessment

After review (discussed in the narrative of a report), the NIA develops a prior, or initial assessment, as described analytically in Figure 3.1a, which gives probability densities for Harry's levels of motivation (*M*), legitimacy (*L*), capability-opportunity (*CO*), and acceptance (*A*) of costs.* The prior for *A* is complex, a mix of a uniform and triangular estimates, because of internal disagreement within the NIA. The prior also includes (Figure 3.1b) a holistic direct estimate of the threat Harry poses (i.e., an intuitive estimate of *T* uninformed by actually using the PFT model). Focusing on Figure 3.1a and recognizing that the curves for motivation and legitimacy are identical, we see that the most likely values ascribed to *M*, *L*, *CO*, and *A* are 4, 4, 5, and 5, but that the uncertainties are large. Overall (Figure 3.1b), the NIA's prior sees it as quite unlikely that Harry is a serious threat (values of *T* between 6 and 10).

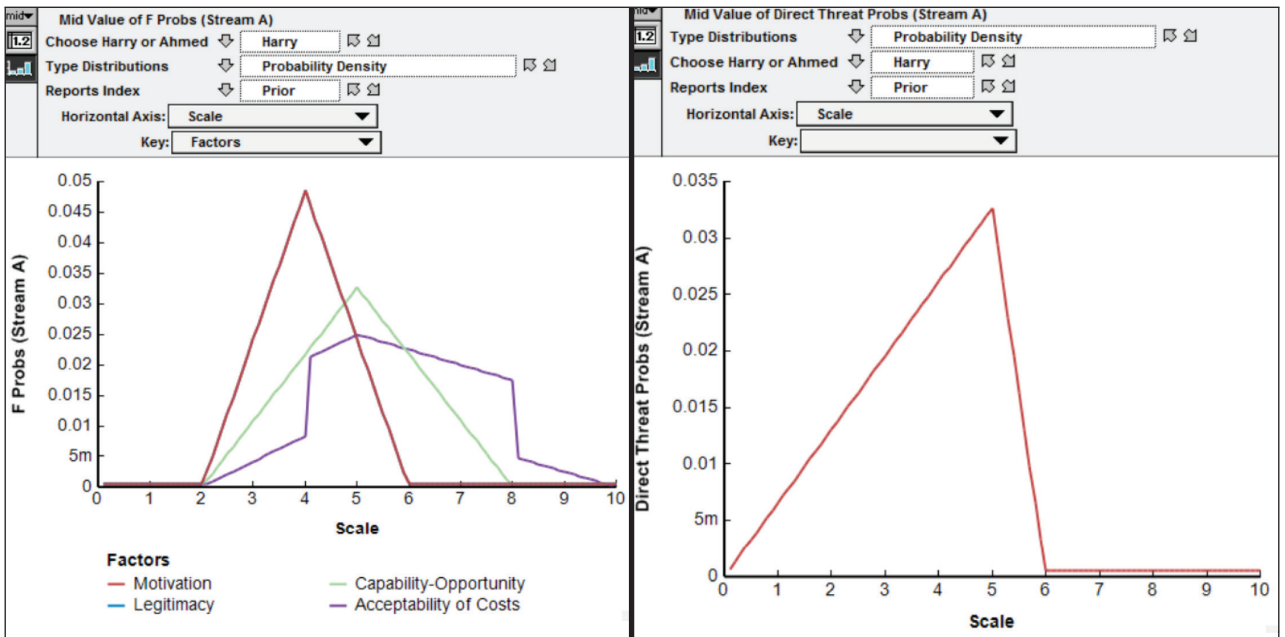
Subsequent Agent Reports

Agents and other sources subsequently provide further raw reports, which are then interpreted by NIA analysts with probabilistic summaries as in Figure 3.1a. Here we merely show excerpts. We do not include the data tables on either estimates or report quality and reliability, but structuring those was an important element of the methodology.

> *Agent A:*
>
> . . . Harry has been working rather steadily for the last two weeks. He landed a job with Eastwood Concrete Company spreading concrete. We have been monitoring his postings on Facebook and other than a few gripes about the way society has treated him, his postings are rather benign. He has opened a Twitter account and . . . has been sharing his views on the ills of society. . . . Although sometimes radical, Harry is articulate in expressing his views. We

* Developing priors at this stage is a crude form of fusion in that it brings together all the information currently available relating to Harry's activities. An appropriate causal model for this fusion process is depicted in Figure 1.3d, the "no-model model relying on input only."

Figure 3.1
Harry Vignette: Priors (Initial Assessments as Probability Distributions)



a. Prior Probability Densities for M, L, CO, A

b. Prior Probability Density for Direct Threat Estimate

NOTE: In panel a on the left, the curves for motivation and legitimacy are identical, and only the red line is visible in the panel.

RAND *RR1200-3.1*

have no reports of him associating with the Slammers in the past two weeks, but that may be because of his full-time job.

Although not in the last two weeks, Harry has been seen repeatedly with die-hard members of the Slammers Club, which is one of the terrorist groups we have been watching closely. I can't see any reason for him to repeatedly socialize with these terrorists unless he's leaning towards becoming one. While I suppose there's a small chance he's that clueless, I kind of doubt it. So that means he's either being spun up to participate in plots—or he's already a conspirator.

Agent B:

Harry talks big and seems to see no problem at all with violence, even terrorist violence, in support of the group's objectives. I have to take that seriously, although he might be just "talking big."

It's hard to tell how motivated Harry is to participate in the Slammers' activity, and from my position in the club membership I do not know if he would be provided with means and tools to participate in a violent act if he should choose to do so. Interestingly, I did hear comments that suggest Harry is nonetheless dubious about paying the cost and taking the risks associated with violent action—even though he has endorsed violence as a legitimate route to right society's wrongs.

From my vantage point, Harry does not appear to be a threat—he likes to talk big about the need for bringing down the system, but never wants to actually do anything other than talk. The few times he's been pushed he always comes up with some excuse about taking care of some pressing need.

Agent C:

I have been watching Harry for some time now, and I don't consider him to be a serious threat, although I am a bit concerned about some of his comments that I overheard at the bar. He said, in effect, that he wished he could do something to support their

cause because he feels it's the right thing to do. It appears that unless he feels it's the right thing to do, he is not interested. However, he confessed that there's not much he can do given his abilities, even though he'd be willing to try given a chance. That sounded to me like an excuse.

Agent D:

I think Harry is a real threat. I have received reports from trusted agents . . . that Harry continues to speak favorably about the Slammers and their violent acts. These agents have befriended Harry and have gained his trust. From what he has been saying lately, Harry is highly motivated—he's just looking for an opportunity to help with an attack. One of my agents observed him training to make explosive devices online. Harry is no innocent naïf—he's a real dangerous character.

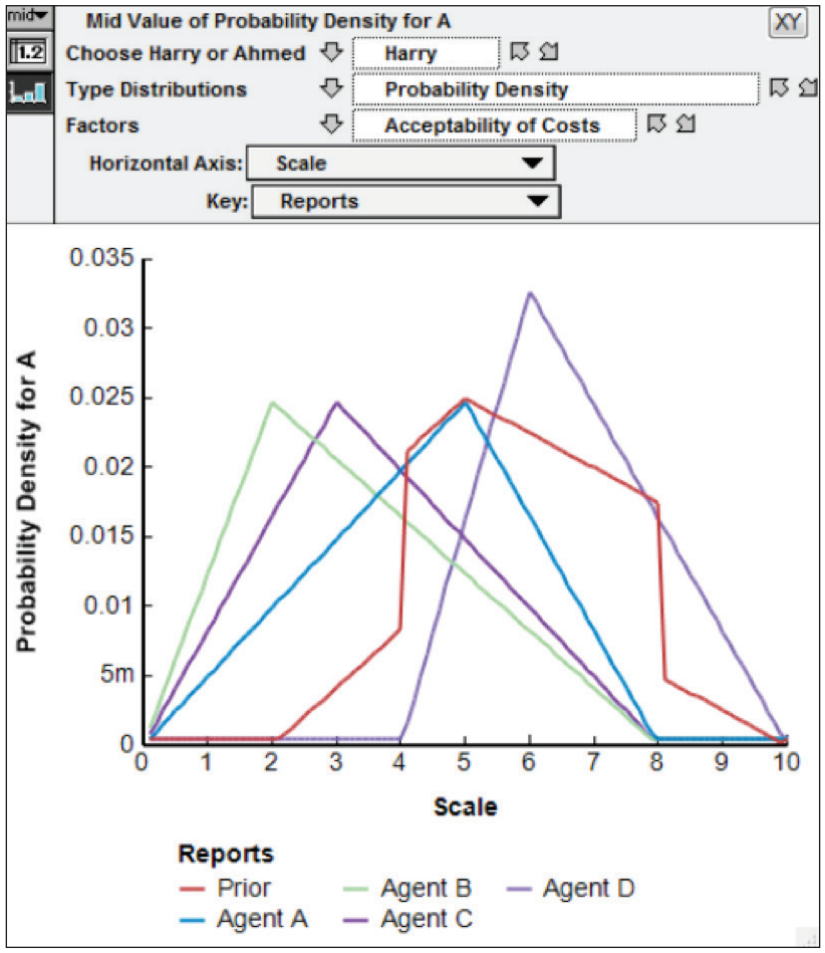
Upon skimming these different reports, it should be evident that the sources are providing very different images. Figure 3.2 illustrates these by comparing across reports for the estimates of acceptability of costs (*A*). If the discrepancies were simply due to chronology or the differences in access of the reporting agents, it might be easy to sort things out, but the sources have different face credibility, the details of their information have different salience, and—in a pinch—the NIA may believe some of the sources are more reliable than others, even if the reports appear equally credible. Such distinctions are reflected not in the reports (as shown in Figure 3.2), but in separate data tables.

Ahmed al-Hiry and the al-Hasqua Group

Excerpt from the initial narrative:

Ahmed al-Hiry is a native U.S. citizen living in Los Angeles with his family. He is 24 years old and is a regular at the 10th Street Mosque. Growing up in Southern California, Ahmed adjusted to the more laid-back California attitudes toward life. He was frequently seen with surfboard in tow catching the waves at Venice

Figure 3.2
Conflicting Assessments About Harry



RAND *RR1200-3.2*

Beach near his home. Other than being forced to attend Friday night services at the Mosque by his parents, he paid little attention to his religion. He had several non-Muslim friends, and he was popular in high school, where he regularly attended school dances in defiance of his parents' wishes.

He was successful at school, which won him a scholarship to study social sciences at UC Berkeley. In college, he participated in sports, joined a fraternity, and generally fit in well with the student population. His grades were good and he graduated on time. However, after graduating, things began to deteriorate. . . . He was . . . forced into partial unemployment, working at menial jobs. To make ends meet, he moved in with his parents again and has been living there ever since.

In college, he injured his left knee, which greatly curtailed his active lifestyle. With nothing to do most of the day, and partially in response to his parents' nagging, Ahmed started attending meetings of Islamist groups in the city. Most have been social organizations aiming merely to improve the lives of the poorer members of the Muslim community. However, one of these groups, the al-Hasqua jihad movement, has a more militant agenda, with some evidence that the group had links to al-Qaeda. This group consists of activists with the secret intent to attack buildings and people in the U.S. that are affiliated with capitalism.

. . . As Ahmed's attempts to find employment continued to fail, he suspected that many employers were prejudiced against him because of his ethnicity. . . . He gave up on trying to find work and began to focus . . . on learning more about what it means to be a Muslim. He abandoned most of his non-Muslim friends, and recently, he has been spending much more time with the al-Hasqua group.

Lately, Ahmed's postings on the group's website and on his Facebook page show . . . that he may be buying into the group's violent agenda. His postings suggest to the NIA that he is affiliated with the group in some way, since their anticapitalist agenda fits with his own feelings of frustration at not finding a decent job despite having a degree. NIA analysts are concerned that, the way he is going, he may intend to commit a hostile act on his own or in participation with al-Hasqua.

The NIA's Prior Assessment

The NIA's prior on Ahmed is in the medium range, but with a great deal of uncertainty. It could be that Ahmed is just disgruntled and enjoys the bravado and companionship of like-minded individuals, but would not think of participating in any attack.

Subsequent Agent Reports

Agent R:

I think it is highly likely that Ahmed has bought into the al-Hasqua cause. . . .

Abu (not an agent, but a friend of Ahmed who became aware of the NIA investigation and offered an impassioned report, with supplementary testimony from others. Abu served in the Army and holds top-level security clearances)*:*

Ahmed is an upstanding American citizen who would never harm this country in any way. I know Ahmed is upset about not being able to find suitable work, but not to the point where he would act against U.S. institutions of any kind—nor would he join a terrorist group like al-Hasqua. He has become very active in the Muslim community of late, but for purely social reasons—not to plot against the U.S. There is no way Ahmed is committed to al-Hasqua, and it's impossible that he joined the group. You guys have it wrong. I don't know where you are getting your information, but I know that Ahmed is a loyal American and would never seriously consider what you suspect him of. I'm including statements by other friends of Ahmed that testify to his integrity and patriotism.

Agent A offers an ambivalent report:

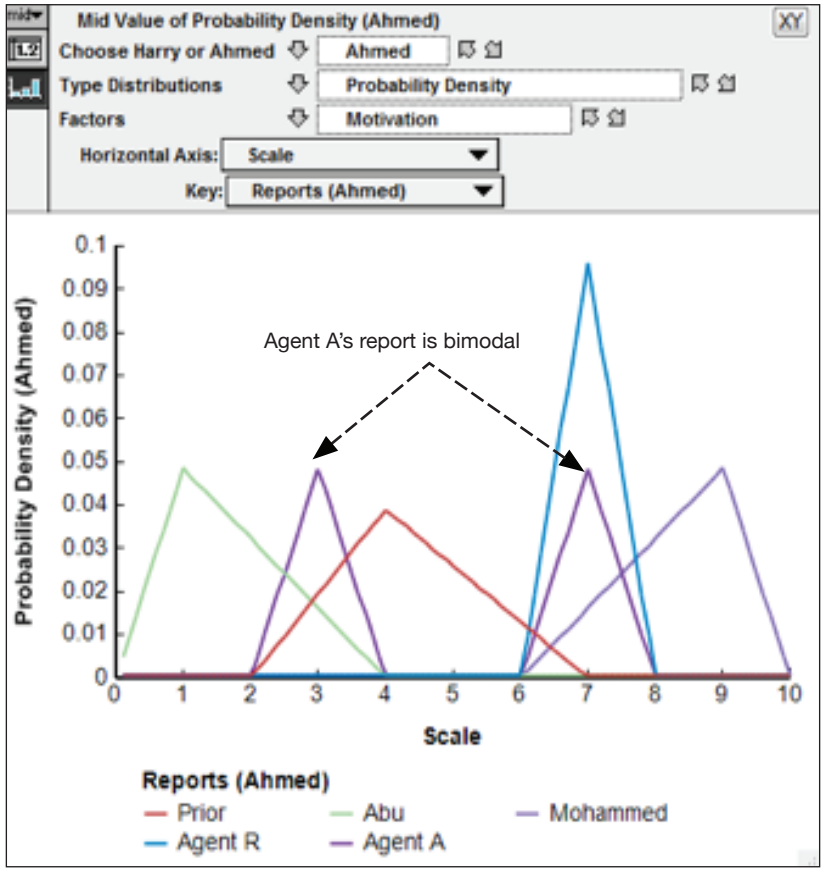
Agent A:

Either Ahmed is an innocent naïf being led astray by members of al-Hasqua or he is a very deceptive true believer. It's hard to tell which. On the one hand, in the community he limits his com-

> plaints about his fate to concerns about unemployment brought on by discrimination as a Muslim. The only action he discusses is lodging complaints. . . . On the other hand, he has been attending al-Hasqua meetings, and his online postings have been a bit more vitriolic. Both when attending meetings and writing online, he uses an alias: Basra al-Noury. I'm not sure however, that he really means what he is posting—he may be just looking for attention given interactions in the wider community.
>
> *Mohammed* (a member of al-Hasqua successfully "turned" by government agents)*:*
>
> For the past two weeks, Ahmed al-Hiry has been attending our . . . meetings. Previously, . . . he was always quiet except when communicating one-on-one. . . . (Now) he has been a more active participant in . . . planning sessions and training curricula. He seems eager to learn all he can about how terror attacks are carried out and what skills one needs to successfully pull off such an attack.
>
> He has also been eager to share his views on the fundamental principles underpinning the al-Hasqua cause. He is constantly accusing local businesses of being biased against hiring Muslims. . . . More and more he is becoming convinced that attacking a few . . . companies will teach the rest a lesson. . . .
>
> I think Ahmed is ready to act against his perceived enemies. He is constantly whipping up the other members and he is rapidly becoming more radical than most group members. . . .

This message stands in stark contrast to that of Ahmed's friend Abu. It could be argued that Mohammed has better information than the others *or*, e.g., that he is embellishing and perhaps even fabricating to provide the information that he believes the NIA wants. Figure 3.3 illustrates the discrepancies for the factor of motivation (*M*). Note that the report from Agent A is bimodal, with peaks at 3 and 7, representing the uncertainty in the narrative above.

Figure 3.3
Conflicting Assessments About Ahmed



RAND *RR1200-3.3*

CHAPTER FOUR

Simple and Bayesian Fusion Methods

This chapter and the next highlight four subjective fusion methods and some variants (1) direct subjective, (2) algebraic, (3) quasi-Bayesian (an approximate Bayesian), and (4) maximum entropy/minimum penalty (MEMP) methods. These seemed particularly suitable for our context. Most of these are useful in combining causal factors to estimate threat *and* for fusing across reports to improve either threat estimates or estimates of the causal factors. The Bayesian method is only for fusing across reports.

In all cases, we evaluate threat by combining across factors with the causal model in question (e.g., the Propensity for Terrorism Model, PFT, described in Chapter One). However, we may do so for each report and then fuse those threat estimates, or we may fuse across reports to better estimate the model factors and then combine.

We considered additional fusion method classes (see Khaleghi et al. [2013] for a survey of fusion methods), of which the most important is Bayesian networks (the term was introduced in the mid-1980s by Judea Pearl, who won the Turing Award in 2012). We deferred applying Bayesian networks because of their complexity, but look forward to revisiting them in future work. Some Bayesian-net studies have already addressed heterogeneous data, uncertainties, and limited data.*

* We did not attempt a full literature review, but two examples illustrate the potential for Bayesian networks in the messy kinds of problems that we are concerned with. One examines uncertainty considerations in military command and control (Das, 1999); another discusses soft factors, including "emotional coherence" in understanding the jury decision in the O.J. Simpson murder trial (Thagard, 2003). Many similar applications exist. As the

In any case, we benefited substantially from the rich literature, especially Pearl's book on causal Bayesian networks (Pearl, 2009).*

The rest of this chapter discusses the direct subjective, algebraic, and Bayesian fusion methods. Chapter Five is devoted to the maximum entropy/minimum penalty (MEMP) fusion method.

Direct Subjective Methods

When combining factors to estimate threat or fusing information across reports, the real-world default is subjective: Those responsible review the evidence without the benefit of analytic aids and express their conclusion directly, although perhaps probabilistically, as in "the odds of the individual representing a serious threat are about one in three." Unfortunately, the human mind does not combine disparate information well, as has long been noted by psychologists in the "heuristics and biases" school.† Thus, it may be possible to do better with analytic aids. As a minimum, it should be possible to add structure and rigor by making assumptions and logic explicit. As will become clear, even if analytic methods do indeed help, subjective assumptions will still play a central role.

Simpson example illustrates, they require assumptions or knowledge that are not usually simple to elicit or represent.

* See also a book demonstrating the practical usability of Bayesian networks in diverse fields (Pourret, Naim, and Marcot, 2008). Another accessible book describes applications to risk analysis (Fenton and Neil, 2012), which includes a foreword by Pearl and an early chapter on the value of causal modeling (pp. 31–50). For unusual Bayesian-net applications involving use of qualitative knowledge, see Wright and Schrag (2014) and Laskey et al. (2004).

† Good introductions exist to this enormous literature (Kahneman, 2002, 2011), and some readers collect notable papers (Kahneman, Slovic, and Tversky, 1982; Goldstein and Hogarth, 1997).

Subjective Algebraic Methods

A substantial relevant literature discusses combining expert-system forecasts and judgments.* One entry point is Clemen and Winkler (2007), a chapter in Edwards, Miles, and von Winterfeldt (2007), or the earlier Clemen (1989).† Other surveys focus more on aggregation of judgment in public-policy and social-choice contexts (List and Puppe, 2009; List and Polak, 2010). An interdisciplinary review at Wharton (Armstrong, 2002, p. 417) drew many of the same conclusions that we did, summarizing with the admonition that one should‡

> combine forecasts derived from methods that differ substantially and draw from different sources. . . . Use formal procedures. . . . . An equal-weights rule offers a reasonable starting point. . . . . Use different weights if you have good domain knowledge or information on which method should be most accurate. Combining forecasts is especially useful when you are uncertain.

This is consistent with our approach as described with the idealization in Chapter Two. The literature has rather consistently concluded, based

---

* Much of the aggregation-of-expert forecast literature deals with problems different from ours in several important respects. The experts are not fabricating information or trying to be deceptive, as occurs with human sources in intelligence and law enforcement. The U.S. government's 2002 assessment about Saddam Hussein's weapons of mass destruction, for example, was unduly affected by the lies of the now-infamous "Curveball" source (Rovner, 2011). Another difference is that the experts are ordinarily being presented with legitimate information, rather than a mix as in Chapter Three.

† An earlier review is Genest and Zidek (1986a, 1986b). The journal volume *Statistical Science*, Vol. 1 (1), 1986, contains a spirited discussion of their paper by several authors and a rejoinder. In particular, Hogarth (1986) emphasized the need to exploit context dependence and to recognize that "expertise" is actually compartmented, comments quite relevant to our current research.

‡ The Intelligence Advanced Projects Agency has an ambitious "Good Judgment" project to improve forecasting for the U.S. government. One of us (Manheim) has been an active participant. The work has been quite innovative, but the questions asked are not as structurally complex as those we deal with, and, as noted in the footnote above, there are other basic differences. For extensive links to the program's publications, see the Intelligence Advanced Research Projects Activity's web page on the topic; as of November 19, 2015: http://www.iarpa.gov/index.php/research-programs/ace.

on empirical evidence, that "simple" methods tend to do better, a conclusion reached in a famous 1979 paper by Dawes, and further corroborated in later work as noted in the Clemen and Winkler review (2007, pp. 162, 174). The question then becomes *which* simple model(s) to use. We adopted three subjective algebraic methods: (1) linear weighted sums (LWS), (2) thresholded linear weighted sums (TLWS), and primary factors (PF). The latter two are relatively intuitive and practical ways to deal with some inherently nonlinear effects. In what follows, we describe the methods for the context of "combining" factors to estimate threat using the PFT model, but the same methods apply when fusing report-by-report threat estimates to arrive at a final threat estimate.

Linear Weighted and Thresholded Linear Weighted Sums

If threat T is calculated by combining independent variables, an easy way to do so is with a LWS.* The PFT model has an array F of factors $\{M, L, CO, A\}$.

The deterministic LWS calculation is simply

$$T = W \bullet F = \sum_k W_k F_k , \tag{4.1}$$

where W is the set of weights. Despite its ubiquity, LWS has serious shortcomings because the actual function determining T can be complex and decidedly nonlinear.

A sometimes-better approach was developed in prior RAND work for a deterministic model (Davis and O'Mahony, 2013, pp. 78 ff). This postulates a linearized approximation with stronger and weaker versions. In its stronger form, which we use here, it amounts to assessing T as 0 unless all of the contributing factors are at or above threshold values.† For example, someone with no motivation for a terrorist

* One paper discusses how to aggregate expert forecasts when the judgments are not coherent and experts sometimes abstain (Predd, Osherson, Kulkarni, and Poor, 2008).

† The method has been used for 20 years in RAND portfolio-analysis work where, in allocating resources, it is essential to provide satisfactorily for *each* of a system's critical components (Davis and Dreyer, 2009).

activity poses no threat even if he sees legitimacy to terrorist violence, has capability and opportunity, and is willing to accept the costs of action. An analogue is the familiar belief that murder is almost always accompanied by motive, means, *and* opportunity. As with most rules, exceptions exist.

Using *TH* to denote the array of thresholds for the components of *F*, the TLWS method for estimating threat with a deterministic model is as in Equation 4.2., which merely states that T is a weighted sum but for a factor Q, which is 0 if *any* of the model factors F are below its threshold TH:

$$T = Q(F,TH)(W \bullet F) = Q(F,TH)\sum_k W_k F_k$$

$$Q(F,TH) = 0 \text{ if } Min(F-TH) < 0 \text{ and 1 otherwise, where} \qquad (4.2)$$

$$Min(F-TH) = \begin{cases} 0 \text{ if } F_k - TH_k < 0 \text{ for any } k,\ k \in \{M,L,CO,A\} \\ 1 \text{ otherwise} \end{cases}$$

For our current work, we had to derive probabilistic versions of Equations 4.1 and 4.2. The formal equations look much the same, but with T and F replaced by the probability distributions $\Pr(T)$ and $\Pr(F)$. The result is

$$\Pr(T) = Q(F,TH)\sum_k W_k \Pr(F_k)$$

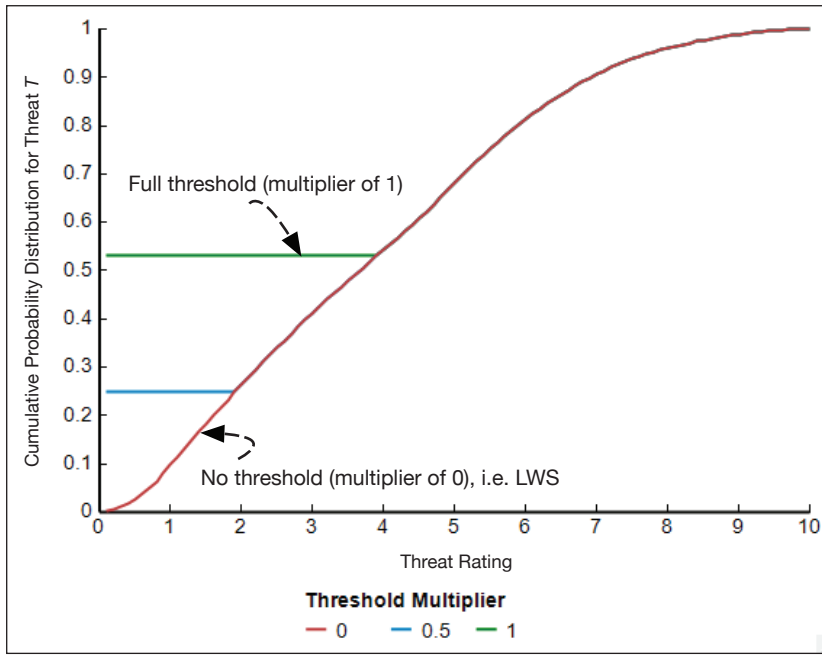
$$Q(F,TH) = 0 \text{ if } Min(F-TH) < 0 \text{ and 1 otherwise, where} \qquad (4.3)$$

$$Min(F-TH) = \begin{cases} 0 \text{ if } F_k - TH_k < 0 \text{ for any } k,\ k \in \{M,L,CO,A\} \\ 1 \text{ otherwise} \end{cases}$$

LWS is just a special case of TLWS in which the thresholds are 0 and Q is 1.

Uncertainty analysis is necessary because TLWS is an approximation and because—even if it is correct—its threshold values are unknown. We facilitate exploratory analysis by inputting default values of the factor thresholds (typically 4) and then allowing the analyst to scale them up or down with a multiplier, as in Figure 4.1, which is for

Figure 4.1
Sensitivity of TLWS to the Threshold



RAND *RR1200-4.1*

an illustrative individual with particular assumed probability distributions for the four factors of the PFT model (motivation, etc.). The plot illustrates a cumulative distribution function for T calculated with TLWS for default threshold values (the top, green line), intermediate thresholds (the blue line), or no threshold (the bottom, red curve).

The differences are substantial: With TLWS and full thresholds (the top line), the individual being rated is said to be more than 50 percent likely to be no threat at all (i.e., to have a T value of 0). In contrast, the LWS calculation ascribes only about 25 percent probability to the individual being in the very low threat category ($T < 2$). The TLWS method is more cautious about labeling someone a threat because the PFT model looks for *all* of the factors to be above thresholds. That may be a good assumption as a baseline, but an analyst might want to know how results would change if he relaxed the threshold requirement. This

is merely one example of how information fusion may be somewhat sensitive to *model* uncertainties (structural uncertainties).

The Primary Factors Method

The primary factors (PF) method is an alternative way to combine across factors to estimate T for a given report (or to fuse report-specific estimates of threat across reports). A deterministic version was developed earlier to represent the phenomenon in which an individual's or group's behavior is dominated by the most important of the several factors that might cause the given behavior (Davis and O'Mahony, 2013, pp. 76 ff). Motivation, for example, can be due to any of a number of causes. In some cases, the primary source of motivation matters far more than the others, although overall motivation might be increased marginally by other factors (e.g., an individual might be highly motivated by the cause alone, but might be a bit further stimulated by enjoying risk and violence). Using the PF method and a deterministic model, the threat T is expressed as a function of the primary and secondary factors P and S by Equation 4.4, a slight change from that in the original source:

$$T = \begin{cases} 0 \text{ if } P + \left(\frac{S}{P}\right)^2 \tau < 0 \\ 10 \text{ if } P + \left(\frac{S}{P}\right)^2 \tau > 10 \\ P + \left(\frac{S}{P}\right)^2 \tau \text{ otherwise} \end{cases} \tag{4.4}$$

where

$$P = Max(M, L, CO, A)$$

and

$$S = Max(F')$$
$$F' = \left\{ F_i' \middle| F_i' \in F, F_i' \neq P \right\}.$$

The parameter τ is an adjustable constant that tunes the magnitude of the secondary factor's effect.) A default value is 2. Results for very small *P* are anomalous but not consequentially so. Similar formulas apply if PF is used to fuse factors across reports.

To deal with probability distributions, we had to extend the method in several ways. To apply the PF operator to a distribution, we represent the distribution by a single metric (typically the mean). We also impose optional threshold requirements for quality and motivation. Finally, at the programming level, we provide functions for going back and forth between deterministic and probabilistic formalisms consistently. If the thresholds have been met, then the result has the form of Equation 4.5 except that Pr(*T*) is on the left and the *P* and variables on the right are replaced by narrow distributions centered at the means of the primary and secondary variables.

The complementarity of TLWS and PF methods can be seen in an example. Suppose that we are estimating threat from the factors of the PFT model. Suppose that the factors' best-estimate values are {4,3,9,5} for *M, L, CO,* and *A*, but with some uncertainties in each case. The TLWS method will return a threat estimate of 5.2 for equal weights and no thresholds. TLWS is doing a kind of averaging. In contrast, the PF method focuses more on the adverse characterizations. Its estimate will be approximately 9.6 with $\tau=2$.

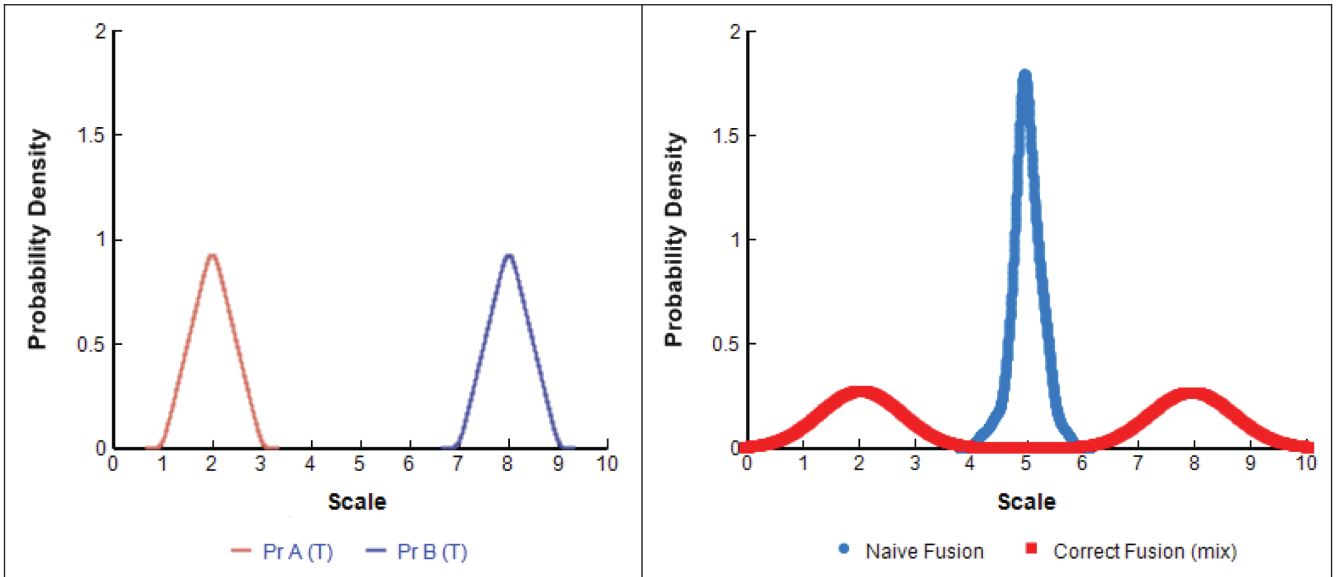
Mathematical Subtleties

Certain subtleties arise in fusion mathematics. Suppose that we have two competing reports with very different probability distributions for *T* (Figure 4.2a). If the reports are equally credible, we might use a probability density that gives equal weight to both, writing this mathematically as

$$\Pr(T) = (1/2)\Pr_A(T) + (1/2)\Pr_B(T). \qquad (4.5)$$

For those comfortable with the language of mathematical statistics, it is straightforward: The fused distribution is to be generated by a mix of the distributions on the right side, with the probabilities of each distribution given by the linear coefficients. To avoid error when using

Figure 4.2
Significance of Correct Mathematics



(a) Competing Reports (probability densities) | (b) Alternative Fused

RAND *RR1200-4.2*

Monte Carlo calculations, however, the analyst must instruct the program to treat the elements of an expression as a mix. Otherwise, the result may be incorrect, because the program does not know to use a mix. For example, if the two individual terms are distributions as in Figure 4.2a, the result for Equation 4.5 might be the unimodal blue distribution in Figure 4.2b, rather than the correct bimodal distribution in red. How to specify the mix depends on the programming language.*

Quasi-Bayesian Fusion

For reasons discussed below, we concluded that we could not use standard Bayesian analysis without information that was unlikely to be available. Thus, we developed a method that we call *quasi-Bayesian* because it *seems* superficially to be Bayesian but is *not really* (hence the label "quasi"). Alternatively, it may just be seen as an approximate version of Bayesian fusion.

The method is largely just a particular application of Bayesian methods with familiar concepts, such as the prior, likelihood function, posterior, and Bayesian-update formula. We discuss those aspects first. We then describe the non-Bayesian way in which we account for differences in reliabilities across reports. It is that which makes the result quasi-Bayesian.

First, let us discuss traditional Bayes updating as a fusion method. The derivation of Bayes updating is described in many texts. Bayes' rule in its elementary form can be expressed as

$$\Pr(H|E) = \frac{\Pr(E|H)\Pr(H)}{\sum_{k \in H} \Pr(E|H_k)\Pr(H_k)} = \frac{\Pr(E|H)\Pr(H)}{\Pr(E)}. \qquad (4.6)$$

* For Analytica, these subtleties are now discussed in Lumina Decision Systems' wiki-based documentation; as of November 19, 2015:
http://wiki.analytica.com/index.php?title=Mixture_distribution

The notation is that H is the discrete hypothesis set; $\Pr(H)$ is the *prior* distribution; and $\Pr(H|E)$ is the *posterior* distribution, the revised estimate after making use of the evidence E. The denominator on the right side is the *probability of the evidence* $\Pr(E)$. $\Pr(E|H)$ is the *likelihood function*; it is the probability of observing the evidence given the particular hypothesis is true.

Application to the Threat-Assessment Problem

In considering a Bayesian method for our problem, we found it useful to make comparisons with simple textbook examples of Bayes' rule, such as when an observer suspects that a coin being flipped successively in a trial is biased and seeks to estimate that bias (see Appendix C). Table 4.1 identifies contrasts between using Bayesian updating for this simple textbook example and using it for the more demanding threat-assessment problem. Reading downward, simple examples of Bayes' rule fully characterize outcome with one attribute, such as whether the face is a head or tail (or what fraction of flips face up heads). In our problem, we have a model with four attributes (M, L, CO, A). Continuing, the simple examples have results driven by chance (a random or stochastic process), while in our problem an individual's attribute values may be constant: It may be only our *knowledge* about the values that is expressed probabilistically. In simple cases, a new piece of evidence is crisp, as in "the coin landed on its head," whereas in our work a report may give a probability distribution for an individual meriting a 1, 3, 5, 7, or 9 on the threat scale. The prior in a simple example may be subjective, but it is uncomplicated. The prior for the threat-detection problem, however, might be the result of a previous study, a mere hunch, or some empirical base rate (but *which* base rate, and how could it be established?). Simple problems assume stationary statistics: For example, each coin flip is independent, with outcome governed by the same mathematical probability. In contrast, an individual's attribute values may not be constant, and, even if they are, observations about them may not be. Simple problems lead to convergence: After a long series of coin flips, we would ordinarily expect updating to converge on a good estimate of the true head bias. In contrast, when fusing counterterrorism reports, estimates may not converge, because of con-

Table 4.1
Contrasts

Issue	Simple Cases	Threat Detection
Number of attributes	One: e.g., the head-bias of a coin	Multiple, as with the factors of the PFT model (*M*, *L*, *CO*, *A*)
Deterministic versus stochastic process	Outcomes are described by a stochastic process	An individual's attributes may be fixed, with unknown values or may change with time. Observation-based reports about them may have random relationships to them.
Certainty of evidence	Observations are crisp, as in result of a coin flip is head or tail	Observations are probabilistic, as with probability distributions for the values of the causal-model factors or the threat level
Prior	Reflects subjective judgment about nature of bias, if any	Might be result of earlier assessment or some empirical base rate. Which base rates are appropriate? How can relevant base rates be obtained or estimated?
Likelihoods	Dictated by well-established mathematics	Subjective and context dependent. Estimates may reflect numerous biases and context dependence may be complex, reflecting all prior knowledge.
Stationarity	Assumed	Uncertain. Actual attributes may fluctuate or change, as may relationship between observations and reality.
Expectation of convergence	Yes	Not necessarily
Significance of updating calculation	Result should converge to correct estimate of coin bias	Very problematic

flicts among the reports, changes in the individual, deception, or other factors.

Most of the issues in the right-most column can be dealt with by somewhat more advanced Bayesian methods; others, not so easily. Overall, we concluded that Bayesian updating will *sometimes* be useful in our fusion efforts and that it made sense to have such a method.*

* The word *sometimes* applies because the information in a new report may not reasonably be construed as an update via the Bayes' formula of the previous assessment. For example, if two sources of information are in major contradiction, the proper inference may be that one or the other is reasonably correct, but not both.

The logical framework can be powerful, even if the inputs are not well known.* To lay out the mathematics, we will illustrate for fusing the report-by-report estimates of the factors of the PFT model. The same method applies for fusing report-by-report estimates of threat to arrive at a final threat estimate.†

If using a model to estimate threat, such as the Propensity for Terrorism (PFT) model described in Appendix B, we have, for each of the factors in that model, multiple hypotheses: $F \in \{M, L, CO, A\} = \{F_1, F_2, F_3, F_4\}$, where each factor can be at one of five threat levels. That is, $F_i = \{1, 3, 5, 7, 9\} = \{\text{very low, low, medium, high, very high}\}$. After estimating Pr*(F)*, we calculate Pr(T) using the PFT model as described earlier. We do that separately using the thresholded linear weighted sum (TLWS) of factors or using the primary factors (PF) method. In both cases, we are exploiting the probabilistic independence of the F factors. For TLWS, this means assuming‡

$$\Pr(T) = \begin{cases} 0 \text{ if any } F_i \text{ is below its threshold} \\ \sum_i W_i \Pr(F_i) \text{ otherwise} \end{cases} \tag{4.7}$$

* The core of Bayesian inference is described differently across sources, but the objects of the formalism (beliefs or credences) are treated as probabilities (Jaynes and Bretthorst, 2003, Chapters 1–2). Estimates should be consistent with known truths, and updating should otherwise spread uncertainty across possibilities as dictated by the likelihood and the prior. "Objective" and "subjective" Bayesians differ philosophically about how to construct priors and even on whether Bayes' updating formula is implied. Most assume that it is, but others argue that instead an update should just reapply the principles of calibration and equivocation (see Williamson, 2010, and a review, Kotzen, 2012). This view sees maximizing information entropy as fundamental, as discussed in Chapter Five.

† We refer to fusing factor-level estimates across reports first, and then estimating threat, as "fuse then combine," and to combining factors to estimate report-by-report threats, and then fusing those across reports, as "combine then fuse."

‡ A Bayesian network representation of the PFT model would define a conditional distribution for T given the four factors. This conditional distribution would simply implement the PFT calculation as a function of the values of the input variables. Here, we use a simpler approach.

For the PF method, we take the means of the factors, check to see that some optional thresholds have been achieved (quality of report and motivation), find the largest and second-largest of the means, and use the PF formula given earlier.

Returning now to fusing across reports to improve factor estimates, Table 4.2 illustrates the input structure for entering four prior distributions, one for each factor of the PFT model (e.g., motivation). A cell value might be (see "Example") the probability that motivation is medium. A similar table would apply if we had priors for the threat estimate.

Applying Bayes' rule in Equation 4.6, we get the following posterior probability for factor F_i:

$$\Pr(F_i|E) = \frac{\Pr(E|F_i)\Pr(F_i)}{\Pr(E)} = \frac{\Pr(E|F_i)\Pr(F_i)}{\sum_{k=1}^{5} \Pr(E|F_k)\Pr_k(F)}. \qquad (4.8)$$

In the threat-detection problem, as in other applications, we may have a very poor prior. We may also have major uncertainty about the likelihood function. That is, the Bayes updating equation is by no means a mere identity as in many applications. Indeed, *Bayesian updating as in Equation 4.8 is describing a new approximation of the posterior probability distribution on the PFT factors based on uncertain updating of an uncertain prior*.

In some cases, then, the "Bayesian" results will be poor—not because Bayes theory is poor but because applying the formula cor-

Table 4.2
Expression of Priors

Level (Value) Model Factors	Very Low (1)	Low (3)	Medium (5)	High (7)	Very High (9)
Motivation			Example		
Legitimacy					
Capability-Opportunity					
Acceptability of costs					

rectly requires information that we lack. In other cases, it may be a reasonable approximation.

Table 4.3 gives an illustrative data table for the likelihood function. The same structure applies to all the factors and to the likelihood function for threat level *T*. The cell labeled "Example" would be the probability that, if the actual level of the variable is 1 (very low), then evidence would report it as level 2 (low, with a value of 3).

The likelihoods (cell values) are substantive, contextual matters. They may also be path-dependent and dependent on relationships between observables and actual attributes. Table 4.4 illustrates the form of evidence. In typical Bayesian work, the evidence from a given

Table 4.3
Data Structure for Likelihood Functions

	Level (Value) Reported				
Level (Value) of Prior	**1 (very low, 1)**	**2 (low, 3)**	**3 (medium, 5)**	**4 (high, 7)**	**5 (very high, 9)**
1 (very low, 1)		Example			
2 (low, 3)					
3 (medium, 5)					
4 (high, 7)					
5 (very high, 9)					

NOTE: The structure applies whether the likelihood function is for a factor, such as motivation, or a threat estimate *T*.

Table 4.4
Expression of Evidence for a Given Factor

Level (Level Value)	Reported Probability
1 (very low, 1)	0.1
2 (low, 3)	0.3
3 (medium, 5)	0.5
4 (high, 7)	0.05
5 (very high, 9)	0.05

report would be a single number, as in "His motivation is high, 7." In our application, a report (an example of evidence) is giving the subjective probability that the variable in question is very low, low, medium, etc. The same structure applies for evidence on threat.

The reason for this complexity is that we expect observations to be equivocal and not definitive.

1. *Observables Versus Factors.* The observations and inferences made by sources may vary randomly around the true values of the factor because observed behaviors are not precisely the same as the factor values. An observer may believe that a statement made by an individual reflects his true beliefs, but it may not. The observer may misinterpret what he hears. Inferences are also subject to context. A person who is actually not motivated for an extremist cause might, in a social setting, make statements of apparent support.
2. *Observer Biases.* An observer's reports may reflect cognitive biases with which a huge literature is concerned (Kahneman, 2011; Goldstein and Hogarth, 1997; Kahneman, Slovic, and Tversky, 1982).
3. *Aggregations.* Where a report is actually an aggregation of inputs from multiple observers, such as an agent's informants, each of whom has biases that may or may not be correlated, the result is a mix. An updated report from the same mix of sources will be some complex updated mix of a previous mix.

Although we don't elaborate on them here, additional complications also arise with multiple reports:

4. *Double Counting and Other Correlations.* If multiple reports reflect the same facts, biases, or estimation approaches, the reports may be redundant rather than independent, creating errors. If human sources tell an investigator "Well, I hear . . ." (but without details), they may all be reporting on what a single individual said or did on a particular occasion. If a report says that a person attends extremist meetings with certain friends,

then another report noting that the person also communicates with those people on social media is not adding information about associations.

It follows that attempting to estimate likelihoods is very challenging, and quite unlike using the mathematics of, say, the coin-flip problem. We propose two approaches, one involving generic likelihoods and one demanding context-dependent estimation of likelihoods.

Generic Likelihood Functions for Exploration

We concluded that the analyst tool kit should include a small number of generic likelihood functions, essentially alternative models that are different in the most important ways. The intent is that viewing results of Bayes updating with the alternative likelihood functions will be informative and give a reasonable sense of the range of possibilities, as in more familiar parametric analysis in physics or engineering.[*] For our prototype work, we considered two illustrative generic likelihood functions. These differ in how *close* we expect observation values to be to actual factor values and whether the likelihood array is *symmetric*.

Model 1: Observations Are Close and Symmetric

Table 4.5a defines a model of likelihoods that assumes observations are probably close to correct about the factor in question and that errors are symmetric. The precise numbers chosen are arbitrary except for assuring the absence of zeros and adherence to the qualitative description.[†] This model, then, corresponds to assuming that observations can

[*] A proper Bayesian likelihood function would be more complicated to establish in our context. A likelihood function at one step of a fusion process should depend on what has been learned from previous reports. That, in turn, would depend on the order in which reports were considered. As an analogy, if one were estimating the head bias of a coin after a series of coin-flipping experiments, the proper likelihood function to use after the first experiment would be different if that experiment had had 1,000 flips rather than five. Our generic likelihood functions are constant and do not reflect a measure of the weightiness of accumulated knowledge.

[†] If a hypothesis has a zero prior, its "Bayesian" update will still be zero despite strong evidence to the contrary. Thus, in using Bayesian methods, one must be careful to allow some non-zero probability for all hypotheses.

Table 4.5a
Likelihood Function if Observations Are Close and Almost Symmetric

Level of Prior	Level Observed				
	1	3	5	7	9
1	0.5	0.35	0.1	0.04	0.01
3	0.35	0.37	0.2	0.04	0.04
5	0.1	0.2	0.4	0.2	0.1
7	0.04	0.04	0.2	0.37	0.35
9	0.01	0.04	0.1	0.35	0.5

Table 4.5b
Likelihood Function if Observations Are Wide and Almost Symmetric

Level of Prior	Level Observed				
	1	3	5	7	9
1	0.35	0.25	0.2	0.15	0.05
3	0.25	0.3	0.2	0.1	0.15
5	0.2	0.2	0.2	0.2	0.2
7	0.15	0.1	0.2	0.3	0.25
9	0.05	0.15	0.2	0.25	0.35

be interpreted straightforwardly. Table 4.5b is a comparable table that assumes that observations will vary more from the actual values.*

Model 2: Observations Are Asymmetric

Table 4.6 shows an alternative likelihood model. This regards it as more likely that an individual that is highly motivated for a terrorist cause will be characterized as having very low or low motivation (bottom left of data field in red) than that a very low-motivation individual will be

* The likelihood functions represented by the tables in this section are used in different contexts and must be interpreted accordingly. The convention is that if one is updating, say from report j to report j+1, then "Level of Prior" refers to the estimate before the update (the estimate as of report j) and the "Level Observed" refers to the evidence from the new report, j+1.

Table 4.6
Likelihood Function if Observations Are Wide and Asymmetric

Level of Prior	Level Observed				
	1	3	5	7	9
1	0.5	0.3	0.15	0.04	0.01
3	0.3	0.5	0.15	0.04	0.01
5	0.025	0.17	0.61	0.17	0.025
7	0.04	0.06	0.1	0.5	0.3
9	0.15	0.15	0.01	0.01	0.68

characterized as being highly motivated (top right). In some contexts, one might expect radical individuals to be deceptive, hiding true feelings from anyone other than closest associates. This likelihood function is representing the kind of reasoning that we associate with a wise and skeptical, or perhaps paranoid, investigator: "Yes, he behaved nicely and spoke moderately in that social setting. So what? That's *exactly* what we would expect an intelligent jihadi to do!" This table anticipates considerable variation between observation and reality.

Other generic likelihood functions are probably necessary as well. For example, analogues to Tables 4.5 and 4.6 (not shown) are symmetric and asymmetric, respectively, but allow for much greater discrepancies between observation and truth. That might be appropriate given, say, a weak prior and little understanding about how close reported observations are able to characterize true attribute values. With the benefit of accumulated knowledge, the analyst might shift to use of a "close" likelihood function.

Context-Dependent Likelihoods

Instead of generic likelihoods, we may use likelihoods estimated for a context. Again, we offer two examples. The first version has analysts confront the prior information and discuss how it should affect their updated reasoning, but then write down their new estimate directly. The second version has analysts develop a likelihood function at the time of their update. To illustrate both versions, assume the prior dis-

tribution described in Table 4.7.* The prior says, in essence, that motivation appears very low or low, but the other factors appear medium.

Direct Updating with "Consideration" of Prior

The team might reason about new information as follows, attempting to rationalize relationships between it and the prior:

> We have no new information on motivation or capability-opportunity, but our agent saw him at a meeting where he made statements suggesting that he either takes the need for serious violence as a given, thinks that it is downright legitimate, or both.
>
> The same agent, however, heard comments that suggest that Harry is dubious about paying the cost and taking the risks associated with the violent action.
>
> Is this information consistent with our previous assessment, or does it suggest changes?
>
> We should update by (a) using the prior for *M* and *CO*; (b) recognizing that his apparent attitude toward *L* might be due to an

Table 4.7
Prior Probability Distributions for Factors

	Level (by characteristic value)				
Factor	**1**	**3**	**5**	**7**	**9**
Motivation	0.4	0.4	0.1	0.075	0.025
Legitimacy	0.05	0.1	0.7	0.1	0.05
Capability-Opportunity	0.05	0.1	0.7	0.1	0.05
Acceptability of costs	0.05	0.1	0.7	0.1	0.05

* This is in the same spirit as "Analysis of Competing Hypotheses" (ACH) in intelligence tradecraft (Heuer and Pherson, 2014), which builds on Heuer (1999), a path-breaking effort to have intelligence methods account for lessons from psychological research. Heuer (2005) is a short discussion of how ACH improves intelligence analysis.

> actual medium-strength belief plus bravado or might be just as the agent inferred, high or very high (the other interpretations are implausible and it seems that the actual levels of medium, high, and very high are about equally probable); (c) updating our assessment about acceptability of costs to consider low and medium values about equally plausible: What the agent heard could reflect modest expression of caution from someone willing, when the time comes, to take risky actions, or it could be literally accurate. Other interpretations appear implausible.

This would be *informed* by the prior (the analysts thought about it), but with no formal Bayes calculations. We have marked the changed items in red in Table 4.8.

Constructing a Contextual Likelihood Function

Suppose instead that the analyst team constructed a likelihood function. The team would want to compare the new information with the prior and worry about how seriously to take both. They might conclude that results are totally inconsistent and that they should pay attention to only one. Or they might instead conclude that the apparent inconsistencies are understandable and even to be expected. They might then want to do Bayesian updating after constructing a likelihood function.

Table 4.8
Direct Updating with No or Minimal Consideration of Prior (No Calculations)

	Level (by characteristic value)				
Factor	**1**	**3**	**5**	**7**	**9**
Motivation	0.4*	0.4*	0.1*	0.075*	0.025*
Legitimacy	**0.002**	**0.002**	**0.33**	**0.34**	**0.33**
Capability-Opportunity	0.05*	0.1*	0.7*	0.1*	0.05*
Acceptability of costs	**0.067**	**0.4**	**0.4**	**0.067**	**0.067**

* Unchanged from prior.

Given the template for a likelihood table above, they might reason as follows:

> Let's talk first about legitimacy. We thought previously that Harry had a medium sense of legitimacy based on earlier partial evidence. Was that wrong? While the previous information is suggestive, this report is reliable, but limited in its certainty as to the true values. We expect that someone with a low or very low value of legitimacy will almost surely be reported as such, with perhaps a little fluctuation. Someone with a medium sense of legitimacy, however, might—in the social context we're seeing—very well come across as more strident in a report than they are in truth. That's perhaps what happened. So, let's just assume that there is some consistency in behavior and ask what we think the likelihood function should be.
>
> [They might claim, perhaps wrongly] Someone with high or very low sense of legitimacy would almost surely be observed that way, or perhaps a bit more convinced. Probably the same asymmetry would exist for the other factors. Anyway, let's assume that is the case and fill out the table that way. These likelihoods are too hard to think about.

The result might be something like Table 4.9 after the analysts assure that likelihoods add to 1 horizontally (they don't need to add to 1 vertically) and (per instruction on methods) assign at least some probability to all possibilities. This table would apply to all the factors. If the table's likelihoods were used in the Bayesian apparatus, the results would be as in Table 4.10. The analyst would want to compare these results with those for the directly estimated new distribution in Table 4.8 (the values are shown in parentheses) and with results from using generic likelihood tables and for non-Bayesian fusion mechanisms. Ideally, this would lead to reflection and iteration. For example, in the imagined ad hoc updating reflected in Table 4.9, the hypothetical analyst team did *not* think about the possibility that a seriously threatening jihadi might be deceptive. The team might reconsider.

Is it worth the trouble to construct such an ad hoc likelihood table? It seems that it would often be worthwhile *if* the analysts were

Table 4.9
Ad Hoc Assessment of Likelihoods

Level of Prior	Level Observed in New Report				
	1	3	5	7	9
1	0.9	0.09	0.0033	0.0033	0.0033
3	0.086	0.82	0.06	0.02	0.02
5	0.01	0.09	0.8	0.068	0.04
7	0.003	0.009	0.097	0.73	0.17
9	0.001	0.001	0.04	0.2	0.76

Table 4.10
Computed Updated Probabilities (direct updates in parentheses)

	1	3	5	7	9
Motivation*	0.4	0.4	0.1	0.075	0.025
Legitimacy	0.009 (0.002)	0.0622 (0.002)	0.338 (0.33)	0.252 (0.34)	0.292 (0.33)
Capability-Opportunity*	0.05*	0.1*	0.7*	0.1*	0.05*
Acceptability of costs	0.05 (.067)	0.447 (0.4)	0.487 (0.4)	0.009 (0.067)	0.007 (0.067)

NOTES: Only the distributions for legitimacy and acceptability of costs were updated. Numbers may not add to 1 due to rounding error. At this stage, the combining methods discussed earlier would be used to calculate an overall threat level.

suitably trained. We say that because people are not generally good at balancing new and old information rationally. However, in a situation where the *real* issue is not so much how to reconcile old and new estimates mathematically, but whom to believe (i.e., which agent has the better track record, which class of digital historical information has proven more reliable over time, which data come from the most "inside" sources, which behavioral sensor seems most trustworthy), then the complications of "elegant" Bayesian or otherwise mathematical fusion will be far less justified and perhaps counterproductive. After all, if the information is contradictory, the best method is not obviously

to do the mathematics assuming contradictions, but rather to focus on the most plausible report(s), as discussed in the next section.

As noted earlier, these examples have been at the factor level, but the structures are the same when working at the threat level. If we were doing a Bayesian update of threat using a new report, and if each report indicates a distinct level for threat, then the analogue to Equation 4.8 would be

$$\Pr(T_i \mid E_j) = \frac{\Pr(E_j \mid T_i)\Pr(T_i)}{\Pr(E)} = \frac{\Pr(E_j \mid T_i)\Pr(T_i)}{\sum_{k=1}^{5} \Pr(E_j \mid T_k)\Pr_k(T)}, \quad (4.9)$$

where the item on the left is the posterior estimate of the probability of threat-level i given the evidence E_j from the new report j. The likelihood functions $\Pr(E_j|T_i)$ is the probability that that the new report will indicate a threat level T_i if the true threat level is T_k. For example, it might give the probability that a report claims that the threat level is low (3) when the true value is high (7).

Reflecting Equivocal Evidence

In our application, the evidence does not always appear as a crisp number, such as Threat = High. Rather, the evidence on factors, such as motivation, and also threat, appears in the form of probability distributions, as described in Chapter Two and Table 4.4. This requires an additional modification when fusing across reports to improve estimates of the probabilities for factors F_i in the array F. Because the factors of the PFT model are nominally assumed to be independent, when we use the Bayesian method to fuse across reports to generate better estimates of the factors (e.g., motivation), we assume that the Bayes formula applies separately to the updates of each factor.

Recall from Equation 4.8 that

$$\Pr(F_i|E) = \frac{\Pr(E|F_i)\Pr(F_i)}{\Pr(E)} = \frac{\Pr(E|F_i)\Pr(F_i)}{\sum_{k=1}^{5} \Pr(E|F_k)\Pr_k(F)} . \quad (4.8)$$

This applies if E is a precise result, such as $F_i = 3$ (Low), or if F_i is binary. If F_i and E are probability distributions, then the notation and concept become more complex. We need to refer to the particular level of the particular factor. Let i be the index for factor, q the index for factor level, and k the index for report. Because the mathematics requires summing over these, we will refer to *i2*, *q2*, and *k2* when necessary. $\Pr(F_{i,q})$ now denotes the probability that factor i has value q.

Suppose that we had just a prior and one additional report, thus with a single update. We then *assume* that the update is the weighted sum of the posteriors that would be generated if the evidence were crisp.* The weighting factors are just the probabilities given in the evidence to each of the levels. Temporarily denoting those with W's for clarity, we then have

$$\mathbf{Pr}(F_{i,q}|E_k) = \sum_{q2=1}^{5} W_{i,q2}\,\mathbf{Pr}(F_{i,q}|E_{q2})\Delta(q,q2) . \quad (4.10)$$

That is, we calculate hypothetical posteriors *as though* the evidence were that factor i is very low, then as though it were low, etc. We then calculate the posterior as the weighted sum, using the actual evidence (probabilities) as the weighting factors. Each term of the summand is a normal Bayesian posterior with crisp evidence. The weighting factors are actually given by

$$W_{i,q,k} = \Pr_k(F_{i,q}),$$

i.e., they are the probabilities given by the kth report (the kth example of evidence) for the ith factor being at level q. If we now assume a series

* This assumption is another way in which our method departs from Bayesian analysis, thereby meriting the prefix "quasi."

of reports, indexed by k, then the update after the kth report is processed is

$$\Pr(F_{i,q,k}\left|E_k\right.)=\sum_{q2=1}^{5} W_{i,q2,k}\Pr(F_{i,q,k}\left|E_{k,q2}\right.)\Delta(q,q2)\,, \quad (4.11)$$

where the $\Delta(q,q2)$ is 1 if $q = q2$ and 0 otherwise and $E_{k,q2}$ is the evidence of report k across the possible levels.

The posterior in the summand is given by Equation 4.12:

$$\Pr(F_{i,q,k}\left|E_{k,q2}\right.)=\frac{\Pr(E_{i,q,k}\left|F_{i,q,k}\right.)\Pr(F_{i,q,k-1})}{\Pr(E_{i,q,k})}=\frac{\Pr(E_{i,q,k}\left|F_{i,q,k}\right.)\Pr(F_{i,q,k-1})}{\sum_{q2}\Pr(E_{i,q2,k}\left|F_{i,q,k}\right.)\Pr(F_{i,q2,k-1})}$$

In most cases, we assume that the likelihood functions are constant, independent of report, in which case they are not indexed by k.

Quality and Reliability: Quasi-Bayesian Analysis

A necessary element of developing what we refer to as quasi-Bayesian analysis was to find ways to reflect report qualities and reliabilities, as discussed in Chapter Three. Ultimately, we postulated a method that takes a linear weighted sum (LWS) of the prior and the tentative Bayesian update. This was motivated by noting that at least some Bayesian analysts assessing data for which some known percentage of data is wrong due to, say, sensor failure, will see the data as emanating from a mixture of sources and may thus construct a weighted sum for the result. The weighting factors would be based on the estimated reliability of the data.*

For our problem, suppose a fusion team has processed a number of reports in the past and has come to the conclusion that its resulting estimate is rather solid and not something to be lightly changed. How,

* A much-referenced paper on this is Duda, Hart, and Nilsson (1976). See also Fenton and Neil (2012, p. 128).

then, might a Bayesian update be accomplished or approximated? Unfortunately, we do not have the luxury of a clean concept, such that a report is either right or wrong. The result is a heuristic approximation.

We prefer to see the issue in terms of "stickiness": When updating an assessment, how closely should the assessment stick to the prior assessment? If the prior assessment is completely reliable, then it should not be changed. If it is not perfectly reliable, but a new report is, then the assessment should be based strictly on the new report. In other cases, the new assessment should be in between. Illustrating this for an updating of threat T as one considers a new report (a similar expression would apply if updating a particular factor estimate with a new report), a linear relationship with an intuitively sensible form is (deferring corrections to assure that the value lies between 0 and 1)

$$\Pr(T\,|\,E_i) = S_i \Pr_{i-1}(T) + (1 - S_i)\frac{\Pr(E_i|T)}{\Pr(E_i)}\Pr_i(T)\,, \qquad (4.13)$$

where S_i is stickiness when updating with the ith report. Stickiness has a value between 0 and 1. The first term corresponds to using the previous assessment, and the second to using the quasi-Bayesian update using the new report's information. E_i stands for evidence presented by the ith report. In the likelihood function in Equation 4.13, $\Pr(E_i|T)$ may be any of the generic alternatives mentioned earlier (e.g., Tables 4.5, 4.6, or 4.9).

Using Table 4.5, for example, "Level Observed" corresponds to evidence, as when someone is reported to be at threat level 5 and "Level of Prior" corresponds to the hypothesis, in this case T. Thus, the likelihood of someone who is actually a medium-level threat (5) being reported as a very low threat (1), would be 0.1.

The question then becomes how to estimate S_i. It should depend on the relative reliabilities of the prior assessment and the new one, but that could be complicated to work with because, in principle, it involves the reliabilities and characteristics of all the previous reports. We had no objective mathematical mechanism for calculating the reliability of an assessment based on updating a previous one. As an initial

heuristic, we settled on the following, which assures that S lies between 0 and 1:

$$S_i = \begin{cases} 1 - \dfrac{R_i}{R_{previous}} \text{ if } R_i \le R_{previous} \\ 0 \text{ otherwise} \end{cases} \tag{4.14}$$

where

$$R_{previous} = Max(R_j \,|\, j \in \{1, \ldots i\}).$$

This departs from Bayesian analysis because it does not "correctly" reflect how likelihood functions should change as evidence accumulates. Nonetheless, it has good attributes:

- It is 1 if the new report is unreliable and stickiness is important; it is 0 if the new report is as reliable as any of the reports received and there is no reason to "cling" to the previous estimate.
- It is small if the new report is almost as reliable as the best of the old ones, thereby "tilting" toward acceptance of the new quasi-Bayesian update.
- It reduces, as it should, to ordinary Bayesian analysis if the reports are all equally reliable and independent (i.e., stickiness is then 0).

We concluded that the heuristic was reasonable, especially given the many other approximations in our work. Over time, a better approximation should be developed that depends on some measure of the "weightiness" of accumulated evidence.*

This said, the heuristic introduces an order dependence to the calculation, which is our primary motivation for referring to it as quasi-Bayesian. If the order dependence observed is consequential, then the analyst may choose to ignore the quasi-Bayesian method as unreliable

* A variant, if one wanted the result to be order-independent, would be to apply Equations 4.13–4.14 to each order of reports and then take an average. We saw it as preferable to think more deeply about the best order of report processing if and when order dependence shows up.

or may go more deeply into the case's specific data. Examining the data might suggest that the quasi-Bayesian method makes sense only with a particular generic likelihood function or with an ad hoc likelihood function that assumes a particular order of processing (see Chapter Six for an example).

CHAPTER FIVE

The Maximum Entropy/Minimum Penalty Fusion Method

Overview

This chapter describes the maximum entropy/minimum penalty (MEMP) fusion method. This method attempts to estimate threat-level probabilities as conservatively as possible given what has been reported and what is assumed. Such an approach protects against certain biases and against overconfidence in assessing the potential threat. To elaborate,

- *Conservative* here has a technical meaning: assuming as little as possible beyond what is known, with "as little as possible" defined in terms of the information entropy of a distribution function.
- *What has been reported* is a combination of facts and assertions from agents, analysts, and sources of tips that bear on the threat posed by an individual. These can include direct assertions of an individual's threat level or assertions about factors bearing on threat. All of these are assumed to be uncertain and may even directly conflict with each other.
- *What is assumed* is a set of rules—i.e., mathematical constraints and parameters that define how we estimate, e.g., the subject's overall threat level given the assertions about risk factors.

The MEMP method accounts for the reliability of the assertions, placing more weight on those that are most credible and salient. It also has no difficulty fusing conflicting assertions. The complexity

of the method grows linearly in the number of assertions rather than exponentially. The method requires solving a nonlinear program, but most formulations will be addressable using off-the-shelf tools such as Microsoft Excel's built-in Solver.*

Mathematically, MEMP finds the estimates of threat that optimize an objective function that is a weighted combination of an information entropy–maximizing term and a penalty-minimizing term, the latter being the mechanism by which uncertain and even conflicting information is expressed. The entropy terms are described in more detail below. The penalty-minimizing term used in this article is the traditional weighted least squares method, which seeks to minimize the square of the deviations between inconsistent assertions about what the estimates should be and the final, fused results. Least squares minimization has a very long history, with the first publication on the technique by Legendre (1805). It has the property of attempting to minimize the total (euclidean) distance between the fused estimate and the inconsistent assertions; additional arguments for using least squares (at least in this report) are described below.

MEMP combines and extends methods from several mathematical disciplines. First, it draws heavily from approaches to maximize the entropy of a discrete probability distribution given a set of hard constraints about that distribution, such as knowing some expected values over the distribution with certainty (Jaynes, 1957a, 1957b; Jaynes and Bretthorst, 2003). For example, Myung, Ramahoorti, and Bailey (1996) use a maximum entropy approach to aggregate the opinions of experts about the moments of a probability distribution, with the analysis running in exponential time in the number of experts.

MEMP draws heavily from a machine learning approach called regularization. Regularized models seek to minimize the weighted sum of a measure that assesses how well a classification or regression predictive model fits the data ("goodness of fit" measure) and a measure diagnosing the entropy of the model (see, for example, Bishop, 2007;

* "Linear growth" means that the complexity of setting up the nonlinear programming grows linearly with the number of assertions, whereas it grows exponentially with Dempster-Shafer Theory or other similar formulations.

Duda, 2004). The use of the entropy measure is intended to prevent the model from overfitting—converging too quickly (i.e., not being sufficiently conservative) to results that reflect the noise in the input data rather than underlying phenomena. Commonly, the entropy measure reflects the number of non-zero parameters in the fit model, as well as their magnitudes, with the goal being to minimize these numbers and magnitudes, reducing the complexity of the fit model. Examples include ridge regression, which in part seeks to minimize the squared euclidean distance of all the parameters (for example, Hoerl and Kennard, 1970); the LASSO (Tibshirani, 1996), or Least Absolute Shrinkage and Selection Operator, which seeks to minimize the absolute value of the fit parameters, ideally driving many of them to zero; and the Elastic Net, which combines the two approaches (Zou and Hastie, 2005). Articles explicitly using a maximum entropy-related measure as the "entropy measure" for modeling across a range of applications include Chiang, Borbat, and Freed (2005), Engl and Landl (1993), Mohammad-Djafari et al. (2002), and Banavar et al. (2010). The least squares measure is commonly used as the goodness-of-fit measure in regularization models, with other options being possible. Mohammad-Djafari et al. (2002) list several, for example, including the Kullback-Leibler divergence for assessing distance between probability distributions (Kullback and Leibler, 1951) and the full range of L-norm distances (with least squares being the L_2 distance).

The MEMP method also accounts for the reliability of the assertions, placing more weight on those that are most credible and salient. It is highly general, modeling both complicated constraints on what is known about threat probabilities and complicated penalties on what different assertions (from field observations or expert opinions) imply about threat probability values. It also has no difficulty fusing directly conflicting assertions. The complexity of the method grows linearly in the number of assertions rather than exponentially. The method requires solving a nonlinear program, but most formulations will be

addressable using off-the-shelf tools, such as Microsoft Excel's built-in Solver.*

It should be noted that MEMP (and the other methods in this paper) have some similarities to expert opinion aggregation mechanisms that generate a fused estimate of an event probability given a number of experts' assessments for that probability. We have already noted Myung, Ramahoorti, and Bailey (1996) using a maximum entropy method; more broadly, a great deal of work has been done in this area. As just a few examples, Satopää et al. (2014) provide a review of aggregation approaches as part of developing a logistic regression approach to aggregation that provides for additional certainty about a prediction if multiple experts are consistently arriving at similar estimates. Predd et al. (2008), building on earlier work by Osherson and Vardi (2006), develop a method that minimizes least squares between the aggregate and experts' predictions, including cases in which the experts' predictions are partially missing or inconsistent, while enforcing that the aggregate predictions are consistent themselves. List and Puppe (2009) provide a survey of methods for aggregating experts' judgments from a logical, social choice theory perspective. That said, while MEMP certainly aggregates these sorts of expert opinions on probabilities, it also is a good bit more general, permitting the inclusion of a wide range of constraints representing known facts about the probabilities and permitting assertions that are more general than the point probability judgments typically seen.

In what follows, we discuss

1. what it means to be conservative
2. how to "model" reports in MEMP, i.e., how to convert reports about a subject into constraints and penalty terms in a nonlinear programming problem

* The complexity of solving the nonlinear programming problem grows as some polynomial function of the number of assertions, depending on the nature of the assertions and the type of nonlinear programming algorithm used, but the result is still computationally easy for any practical data sets in this type of application. (The nonlinear program's complexity growth can indeed be linear in the number of assertions in some situations.)

3. HOW to implement the PFT model as nonlinear programming terms, constraints, and parameters
4. computational considerations.

We illustrate concepts with the Propensity for Terrorism (PFT) model used throughout the report and defined in Appendix B. This refers to overall threat as T and to four input factors: motivation (M), legitimacy (L), capability-opportunity (CO), and acceptability (A) of costs. These are defined on a 0 to 10 scale, but in this chapter we use a discrete version with five tiers (1, 2, 3, 4, 5) having respective characteristic values of 1, 3, 5, 7, and 9. The linear array (vector) of the four factor values is given by

$$\Pr(F) = [\Pr(M), \Pr(L), \Pr(CO), \Pr(A)]. \qquad (5.1)$$

This chapter frequently uses an index over tiers rather than values. The notation $\Pr(M_i)$, then, refers to the discrete probability that motivation's value is that of tier i or $\Pr(M = i)$.

Being Conservative: Maximizing Uncertainty When Assessing What We Really Know

It is possible to generate probability distributions that, in the mathematical sense, imply the least "information" beyond those implied by the constraints. Mathematically, this is a distribution with maximum *information entropy*. The concept of information entropy traces to the 1948 work of Claude Shannon in communication theory (Shannon, 1948; Ash, 1990). The related principle of maximum entropy has had profound influences in decision theory, statistical physics, and other disciplines. Much of this was pioneered by Edwin Jaynes (Jaynes, 1957a, 1957b; Jaynes and Bretthorst, 2003).

To define information entropy for the discrete case in the context of threat fusion, let T be a discrete probabilistic variable representing the overall threat of a subject that takes on values T_i with probabilities $\Pr(T_i)$ as noted above. Then $H(T)$, the information entropy of the distribution for T, is

$$H(T) = -\sum_{i=1}^{5} \Pr(T_i) \log_2(\Pr(T_i)) \,. \tag{5.2}$$

The choice of the logarithm's base is arbitrary. The most common and original choice for use with discrete distributions, however, is to use $\log_2$, which yields information entropy measured in "bits." We assume use of $\log_2$ throughout most of this chapter. Significantly, this form is not arbitrary; it can be *derived* by demanding that the measure called entropy H should do all of the following:

- grow with the number of possibilities for a system's state
- obey the "extensive property," so that if a system C is created by combining statistically independent systems A and B, then $H(C) = H(A) + H(B)$
- be such that H is 0 when only one state is possible
- be such that the $\Pr(T_i)$ are probabilities.*

Ludwig Boltzmann and Robert Gibbs postulated the basics of this concept late in the 19th century in the narrower context of understanding thermodynamic entropy in terms of molecular physics.

The maximum entropy problem, then, is to find a set of probabilities that maximize the above equation, subject to any known constraints on those probabilities. In the extreme case where there are no known constraints (i.e., no information on what the distribution might be), $H(T)$ is maximized when all probabilities have equal values. For N possible values, the value is $1/N$ and $H(X) = \log_2(N)H(X)$. For the five discrete levels used in this report, the probability of each is 0.2 and $H(X) = \log_2 5$, or about 2.33 bits.

Why seek to maximize entropy when estimating a distribution? Any other distribution incorporates additional *assumed* information, the inclusion of which may be unjustified. That is, maximizing entropy assures what Jacob Bernoulli and Pierre Laplace had in mind with concepts referred to as the "principle of indifference" or the "principle of

* These conditions can be found in various places in the literature (Jaynes and Bretthorst, 2003, Chapter 11) and, in the statistical physics literature, Katz (1967, Chapter 2).

insufficient reason" (see discussion in Jaynes and Bretthorst, 2003). The distributions with the maximum possible entropies have equal probabilities, which are appropriate to use when modeling atomic states in cases in which no information about the possible probability values are known.*

Modeling What Has Been Reported About a Subject

Modeling Facts: Introducing Constraints

If there are known facts about the distribution or expectation of T, we can represent them as constraints, leading—in a computational approach—to a nonlinear programming problem. As examples, we may have inequality bounds on probabilities such as knowing that a given $\Pr(T_i)$ is at least greater than and/or less than a particular value. Thus, for a given p_i, we may have $L_i \leq \Pr(T_i) \leq U_i$, where L_i and U_i are lower and upper bounds, respectively, on $\Pr(T_i)$. Both of these bounds are between 0 and 1. Similarly, we may have constraints on the expectation of T, such as

$$E[T] = \sum_{i=1}^{5} T_i \Pr(T_i) \leq \mu_U \ , \tag{5.3}$$

* The indifference principle, namely forced assignment of equal-value distributions, should be applied *only* to atomic states or in cases in which we genuinely have zero information. A common mistake is to pose some problem such as "The statement X might be true or it might be false. Since I have no knowledge on the matter, assume that it is equally likely that X is or is not true." Such reasoning is usually wrong, for at least two reasons. First, depending on the nature of X itself, there may be many more logical ways in which X could be true than false, or vice versa. Suppose that X being true corresponds to an outcome of 7 in the game of craps, in which case X being not true (i.e., ~X) corresponds to "something other than 7"). In that instance, if we do not know the outcome, we should bet heavily on ~X. A second problem is leaping too quickly to the assertion that "we know nothing about this." Suppose that a dimly visible animal could be a horse or a zebra. We *could* assume equal likelihoods, but for those living most places in the world, the no-special-knowledge situation would argue for betting heavily on the animal being a horse. For our context of threat assessment, the states "threat" and "not threat" are not atomic states.

where μ_U is an upper bound on the expectation of T. Finally, we require that the $\Pr(T_i)$ must sum to one, and must all be between 0 and 1, so that the T_i's form a valid probability distribution. An example of a nonlinear programming problem, then, is

$$\begin{aligned} &\text{Minimize } -\sum_{i=1}^{5} \Pr(T_i)\log_2(\Pr(T_i)) \\ &\text{Subject to } L_i \le \Pr(T_i) \le U_i, \forall i \\ &\qquad \mu_L \le \sum_{i=1}^{5} T_i \Pr(T_i) \le \mu_U \\ &\qquad 0 \le \Pr(T_i) \le 1_i, \forall i \\ &\qquad \sum_{i=1}^{5} \Pr(T_i) = 1 \end{aligned} \tag{5.4}$$

As an additional example, suppose that we have evidence suggesting that the probability of a subject's overall threat level being "medium" is less than the total probability of a subject's motivation being either "high" or "very high," $\Pr(T_4)$ and $\Pr(T_5)$. This assertion can be represented as a constraint added to our nonlinear programming problem:

$$\Pr(T_3) \le \Pr(T_4) + \Pr(T_5)\,. \tag{5.5}$$

Statements like this, much less statements on expectations of T, are harder (although not impossible) to address using Bayesian inference. Also, in the Bayesian framework one has to estimate how to distribute the residual uncertainty: Should the probabilities of tiers 4 and 5 be assumed equal, or should some other assignment be made? A substantial and sometimes confusing literature exists attempting to understand the basis for and limitations of related assumptions in the Dempster-Shafer theory. The MEMP method provides one theory-based principle for allocating residual uncertainties across possibilities.

Modeling Assertions: Introducing Penalty Functions

The nonlinear programming examples above assume that any assertions about the distributions are firm and should be reflected as hard constraints on the distributions. In practice, the "facts" are often uncertain assertions supported by varying degrees of evidence. These may contradict each other, as when analyst A is convinced that a subject is a high threat, while analyst B is convinced the same subject poses no threat. In such cases each assertion may provide a "constraint" of some sort on the probability distributions—but what sort, especially when we recognize that the "constraints" can't be satisfied simultaneously?

We deal with such uncertainty by creating *soft constraints.* These do not require the probabilities to be the specified values, but instead we impose *penalties* when the probabilities emerging from the nonlinear program are different from the specified values. We then revise the nonlinear program's objective function so that it attempts to optimize a weighted combination of maximizing the entropy of the threat distribution and minimizing the penalties incurred for deviating from analysts' assessments. As noted above, this is analogous to the machine learning approach of regularization. Regularized models also seek to minimize the weighted sum of a measure that assesses how well a classification or regression predictive model fits the data ("goodness of fit" measure) and a measure diagnosing the entropy of the model. This idea makes intuitive sense, but it adds the complication of having to specify the trade-off between entropy maximization and penalty minimization. Approaches to making these trade-offs are discussed in a subsequent section.

What should the penalty functions be? We see the key desirable properties as

- penalize larger deviations disproportionately more than smaller ones
- scale the penalties to reflect the reports' relative quality and reliability
- express penalties in a way easy to use when performing nonlinear programming: penalty functions should be smooth (no discontinuities) and convex

- use methods with successful history in similar analysis situations of wanting to minimize some overall difference between competing claims over what an output should be.

We found that a penalty function with all of these properties is a weighted quadratic function in which penalties are proportional to the square of the deviation between the asserted value and the fused estimate. As an example, suppose we have an assertion j that the correct estimate for $\Pr(T_i)$ is $\Pr(T_i)^*_j$. Then a penalty function is

$$\phi_j(T_i) = C_j[\Pr(T_i) - \Pr(T_i)^*_j]^2 . \tag{5.6}$$

Here, C_j is a penalty weight. In comparison to the desired properties listed above:

- The value of this penalty function is 0 when the fused estimate of $\Pr(T_i)$ equals the asserted value and increases with the square of the deviation from the fused estimate.
- The value of C_j can be increased as quality and reliability of the assertion increases.
- Quadratic penalty functions are smooth and convex; further, a wide range of solvers exists to solve quadratic optimization problems.
- Minimizing the square of deviations between a fused estimate and conflicting assertions on what the estimates should be is an instance of *minimizing least squares*, the typical approach to identifying approximately optimal solutions to *over determined systems* (i.e., problems in which there are more conflicting assertions about what parameters should be than there are parameters). Least squares minimization has a long history, dating back more than two centuries (Legendre, 1805).

Now suppose we have a series of j assertions, each with different claims $\Pr(T_i)^*_j$ on what values for the probabilities $\Pr(T_i)$ should be. Further suppose that each assertion j has been assessed for reliability,

such that each has been assigned a penalty weight, C_j, with more reliable assertions being assigned larger weights.

Our revised nonlinear programming problem is to minimize a weighted sum of the entropy-related terms and the penalty terms, subject to specified constraints. An example is (expanding on the example presented earlier)

$$\text{Minimize } -\lambda\sum_{i=1}^{5}\Pr(T_i)\log_2(\Pr(T_i))+\sum_{i=1}^{5}\sum_{j=1}^{5}C_j[\Pr(T_j)-\Pr(T_i)^*_j]^2$$

$$\text{Subject to } L_i \le \Pr(T_i) \le U_i, \forall i$$

$$\mu_L \le \sum_{i=1}^{5} T_i \Pr(T_i) \le \mu_U \tag{5.7}$$

$$0 \le \Pr(T_i) \le 1, \forall i$$

$$\sum_{i=1}^{5}\Pr(T_i) = 1.$$

In this formulation, λ is a relative weight on the entropy-related terms, rather than the penalty terms. Setting this weight is discussed in a subsequent section.

Mathematical programming often assumes constraints in the form of equalities, except for a set of baseline inequalities that require the decision variables to be greater than or equal to zero. However, to say that $x \le a$ is equivalent to saying that $x = a - b$, where b is a variable restricted to be greater than zero. Thus, we can work with assertions that are inequalities, namely claim that $\Pr(T_i) \le \Pr(T_i)^*_j$ or that $\Pr(T_i) \ge \Pr(T_i)^*_j$ by adding *slack variables*, s_{ij}, as follows:

$$\Phi_j(T_i) = C_j[\mathbf{Pr}(T_i) - \mathbf{Pr}(T_i)^*_j - s_{ij}]^2, \tag{5.8}$$

where $s_{ij} \ge 0$ for the soft constraint $\Pr(T_i) \ge \Pr(T_i)^*_j$ and $s_{ij} \le 0$ for the soft constraint $\Pr(T_i) \le \Pr(T_i)^*_j$.

An assertion might also assess a condition on multiple probabilities, such as

$$\Pr(T_3) \leq \Pr(T_4) + \Pr(T_5) + 0.1. \tag{5.9}$$

This would be an extension of our earlier example. We can penalize this example as:

$$\Phi_j(T_i) = C_j[\Pr(T_4) + \Pr(T_5) - \Pr(T_3) + 1 - s_{ij}]^2,\ s_{ij} \leq 0. \tag{5.10}$$

More broadly, we can penalize assertions that place an inequality constraint on a function of any of the elements of $\Pr(T)$ as follows:

$$\Phi_j(T) = C_j[f_j(\Pr(T)) - s_{ij}]^2,\ s_{ij} \leq 0. \tag{5.11}$$

Modeling the Reliability of an Assertion

In this section, we discuss methods for setting the size of the weights (C_j's) on the penalty functions. Chapter Two introduced the idea of quality and reliability of a report, with both being measured as numbers between 0 and 1. These can be computed products of other input factors (quality = credibility × salience), or they can be set subjectively, perhaps using the computed products as a decision aid. Intuitively, a report with a score of 1 is considered to be as close to certainty as one can reasonably get when working with security reports; a report with a score of 0 is considered to provide no useful information whatsoever, such that one could discard it with nothing lost.

It is possible to derive a justification for using the simple scores from 0 to 1 as the weights (C_j's) on the penalty functions, on the grounds that doing so is a good approximation for one of the top sophisticated weighting rules out of the field of machine learning. We derived that in a longer version of the report but do not include it here. For the examples in this report, we use the 0–1 scores for the assertions as the C_j's.

Assertions in Time: Conflicting Claims with Differing Time Stamps

We may have situations in which reports about a subject come in over time, and the nature of assertions being made change over time, as well. Our approach to information fusion allows for processing reports

in different orders, because results may depend on order. The default approach is to process reports in the order of their arrival, i.e., on their "time stamp." With some fusion methods (e.g., the Bayesian methods), a new report will "update" the previous one. However, later reports may actually be better or worse than older ones in terms of credibility, salience, and even the subjective tie-breaking assumptions about the relative weights to be given to reports of equal quality. In Chapter Four, we discuss how we deal with such issues using a heuristic.

With the MEMP approach it is possible, even natural, to treat all of the reports simultaneously, regardless of their time stamp. Each report's assertions are represented by constraints in the nonlinear program, even assertions that are contradictory. This can be seen as an advantage of the MEMP method, or at least a feature that distinguishes it from some of the others, notably Bayes updating.

Weighting the Entropy-Maximizing Term

Recall that the objective function of the nonlinear program had a weight λ in front of the entropy-maximizing term. In this section, we discuss mechanisms to assign a value to λ. In machine learning and/or predictive analytics applications, the λ is set to be whatever value led to the greatest predictive accuracy during the model training and testing process, assuming the analyst had a great deal of data linking real-world inputs to real-world outcome values. In our application, we do not have real-world examples to test, say, whether a given subject's assessed value of $\Pr(T_i)$ should be assumed accurate, but a given agency might have such empirical information. We could pursue the matter to some degree with fictitious data or with real-world data on terrorism in the public media.

To get an approximate sense of how the entropy-maximizing terms behave in comparison with the minimum-penalty terms, which are quadratic terms, we compute a second-order Taylor-series expansion for the divergence terms, as follows.

Each set of entropy-maximizing terms has terms such as $\Pr(T_i)\log_2\Pr(T_i)$. The first and second derivatives of the entropy-maximizing terms have terms that look like the following when expressed temporarily in terms of the natural logarithm ln.

$$\frac{\ln \Pr(T_i)+1}{\ln 2} \text{ and } \frac{1}{(\ln 2)\Pr(T_i)}$$

These terms are collectively maximized when all the $\Pr(T_i)$ terms are equal (each equal to $1/N$). Taking the second-order term of a Taylor-series expansion of this means that, around the optimal values, the maximum-entropy term will behave approximately as a quadratic penalty function with weight:

$$\frac{1}{2} \cdot \frac{1}{(\ln 2)(\frac{1}{N})} = \frac{N}{2\ln 2}$$

With five threat levels (as is standard for threat models in this report), this weight will be approximately equal to 3.607.

In practice, we will probably want to weigh the objective of maximizing entropy similarly to weighting deviations from an analyst's report (probably a "highly reliable" report). To perform the weighting, we note the following:

- There are N entropy terms (one for each probability estimate), which means that we need to divide each term by N so that maximizing entropy has the same approximate weight as conforming to an analyst's ratings.
- Assume that we have a desired weight, ω, on maximizing entropy. Using the fact that maximum-entropy terms behave approximately as penalty functions with weight $N/(2\ln 2)$, we can create maximum-entropy terms that behave approximately as if they were weighted by ω simply by multiplying each maximum-entropy term by a factor of $(2\ln 2)/N\omega$.

Combining the above two points means that we should multiply each maximum-entropy term by a factor of

$$\lambda = \frac{\omega \ln 2}{N^2}.$$

As we have emphasized throughout the report, uncertainty analysis is very important in heterogeneous fusion. It should be relatively easy to do systematic sensitivity testing to see how strongly threat estimates made with MEMP depend on the semi-arbitrary assumptions about λ.

Specification for an Illustrative Application Using the PFT Model

Most of the discussion in this chapter has assumed that the incoming assertions are all directly about threat. If instead they are about the factors of the PFT model, then could MEMP be used to find the distributions of those factors that maximize entropy? Also, given assertions from various sources about both the factors and their direct estimates (often intuitive) about threat, could MEMP be used to find distributions for both factors and threat? What would a specification for an MEMP calculation look like? The answer is "It depends."

Initial Architecture: In our basic architecture for this phase of work, MEMP is conceived as an alternative fusion method that can be represented as a function. It can receive inputs at the threat level for each of the reports and return a fused distribution function for Pr(T). Alternatively, it can receive inputs at the factor level for each of the reports and return fused distributions for each of the factors (e.g., M, L, CO, and A). In the latter case, subsequent processing would generate a final estimate of T by combining factors using the TLWS or PF method. In this version, MEMP includes the combining process, whether TLWS or PF, reflecting them in constraints. Also, in this version, MEMP receives inputs of T, M, L, CO, and A by report and returns probability distributions for T. This standalone version of MEMP is convenient to have.

Extended Version: As an alternative (not implemented in the Analytica platform described earlier), the MEMP method can be used to take all available assertions, whether expressed in terms of factors or threat, and return probability distributions for T.

Computational Considerations

Getting to Best: Conditions for Optimizing

Microsoft Excel and Analytica both use nonlinear programming solvers that rely on the Generalized Reduced Gradient (GRG) algorithm (Frontline Systems, Inc., 2015). This algorithm, along with most other nonlinear program solvers, is typically guaranteed to find only a local optimum to a problem. Here, "local" means that the "solution" found is better than all other points immediately around it, but may not be the best solution possible. It is of interest to identify conditions under which a nonlinear solver such as GRG will find the globally optimal solution to an MEMP problem, which is equivalent to asking for conditions under which any local minimum found will be globally optimal, as well.

A key result from the field of *convex optimization* is that if the objective function is a *convex function*, and if the constraints form a *convex set*, then any local minimum discovered will be a *global minimum* (i.e., the genuinely optimal solution) as well. Intuitively, the graph of a convex function looks like an upward-facing bowl, which immediately implies that any local minimum it has must also be a global minimum.

Key functions known to be convex are least squares penalty terms in which we square the difference between the estimated value and a target value, as with our penalty functions. More broadly, squaring the difference between an estimated value and a weighted sum of multiple inputs is also convex. The maximum-entropy terms of the form $\Pr(T_i)\log_2\Pr(T_i)$ are also known convex functions. Finally, the sum of convex functions is also a convex function, meaning that the types of MEMP functions described above, which penalize deviations from single-point assertions of probabilities or expectations, are convex. Thus, in the cases and examples discussed above, the objective function will be convex.

That leaves the question of whether the constraints form a convex set. The set of constraints forms a convex set if, for any two feasible solutions, any point on the line segment between them is also a feasible solution. Suppose we have the following nonlinear program:

$$\begin{array}{ll} \text{Minimize} & f(\mathbf{x}) \\ \text{Subject to} & g_i(\mathbf{x}) \leq 0, i = 1,2, \ldots, m \end{array}$$

Each constraint creates an upper bound. If every function $g_i(\mathbf{x})$ in the constraints is a convex function, then values of that function form a lower bound, and the feasible region of the constraint will then form a convex set. The intersection of the constraints' feasible regions will themselves form a convex set, as desired.

A key result in convex optimization is that linear constraints (simple bounds on values or bounds on weighted sums of values) form a convex set provided that the nonlinear program has feasible solutions. The types of constraints we consider do form convex sets, assuming the constraints permit a feasible solution."

Computational Considerations for Solving MEMP Models

The built-in Solver in Excel has limits of 200 decision variables and 100 constraints (Frontline Systems, Inc., 2015). Our PFT-based MEMP models will have 25 decision variables, for five risk-level probabilities for overall threat plus four input risk factors. There will be at most a few dozen constraints, based on what variant of functional relationship between overall threat and the input risk factors is used. The penalty terms on the analysts' assertions are all simply added to the objective function, so as a result there is no hard limit on the number of such terms. Thus, MEMP nonlinear programs should fit comfortably into the Excel or Analytica solvers.

Next Steps for MEMP

The primary next step for MEMP research is to use it with realistic data. This could be done with a sample of cases including both positives (actual U.S. terror plots) and negatives (cases in which a subject was heavily investigated or arrested but later found to be not a terrorist). Reasonably decent data on such matters are available in the open media. See, for example, Strom et al. (2010), which provides an assessment of how U.S. terror plots have been foiled—and what went wrong

when they did not—based primarily on open-source analysis. Part of the research would involve empirical modeling to provide insight on parameters (most notably, the weight lambda on the maximum-entropy terms, discussed above). More important would be using MEMP in a simulated analysis environment, seeing what happens as we use MEMP to work through the reports that come in about a future terrorist (or nonterrorist wrongly accused) in real life, as well as how MEMP might have been used to make more effective decisions.

CHAPTER SIX

Illustrative Analysis

The analytic platform for experimentation came together toward the end of our research. We were able to do enough experimentation to demonstrate many of the concepts and to verify that the desired flexibility had been achieved. This chapter provides illustrative results, but it will take future work to assimilate what we have (i.e., to gain experience through experimentation) and to refine the methods and platform. The results shown were not crafted for drama, but rather generated by applying the methods straightforwardly to the synthetic data of Chapter Three.

A Flexible Suite of Fusion Methods

Some Illustrative "Standard" but Uncertainty-Sensitive Results

Chapter Two's Figure 2.1 described our conceptual architecture for an analytic platform to enable heterogeneous fusion. The prototype implemented that architecture. Figure 6.1 shows illustrative results using the synthetic data from Chapter Three. It needs explanation because it incorporates considerable information. Each bar shows the assessed probability, after fusion, that the individual in question should be classified as a threat, in the gray area, or as a nonthreat. These categories represent an aggregation corresponding to threat levels T of 6 to 10, 4 to 6, and 0 to 4, respectively. Each clump of bars shows how results change as a function of the fusion method used, as indicated by bar colors. The methods are primary factors (PF), thresholded linear weighted sums (TLWS), Bayesian (called quasi-Bayesian in text), and

maximum entropy/minimum penalty (MEMP). The dashed horizontal lines indicate the corresponding initial assessments (i.e., the "priors"), before the fusion analysis over subsequent reports.

At the top of Figure 6.1 we see a set of "Slicer Bars" (using Analytica terminology). These show the values of selected other variables and allow the user to see effects of changing them. Each has a menu, indicated by the arrows. The analyst can explore how results change as all of these assumptions change. Navigation through the outcome space is nearly instantaneous. This illustrates how uncertainty analysis can be accomplished routinely, rather than being constantly deferred.

Reading down the list of slicer bars, Figure 6.1 shows results for Chapter Three's Harry vignettes. It shows results from only Stream A of analysis. Fusing takes the fuse-first approach, in which factors such as motivation are evaluated by fusing across reports, after which threat is estimated by combining the factors with the TLWS method. For the results based on quasi-Bayesian methods, the asymmetric likelihood function is used. The reports are processed in the order received.

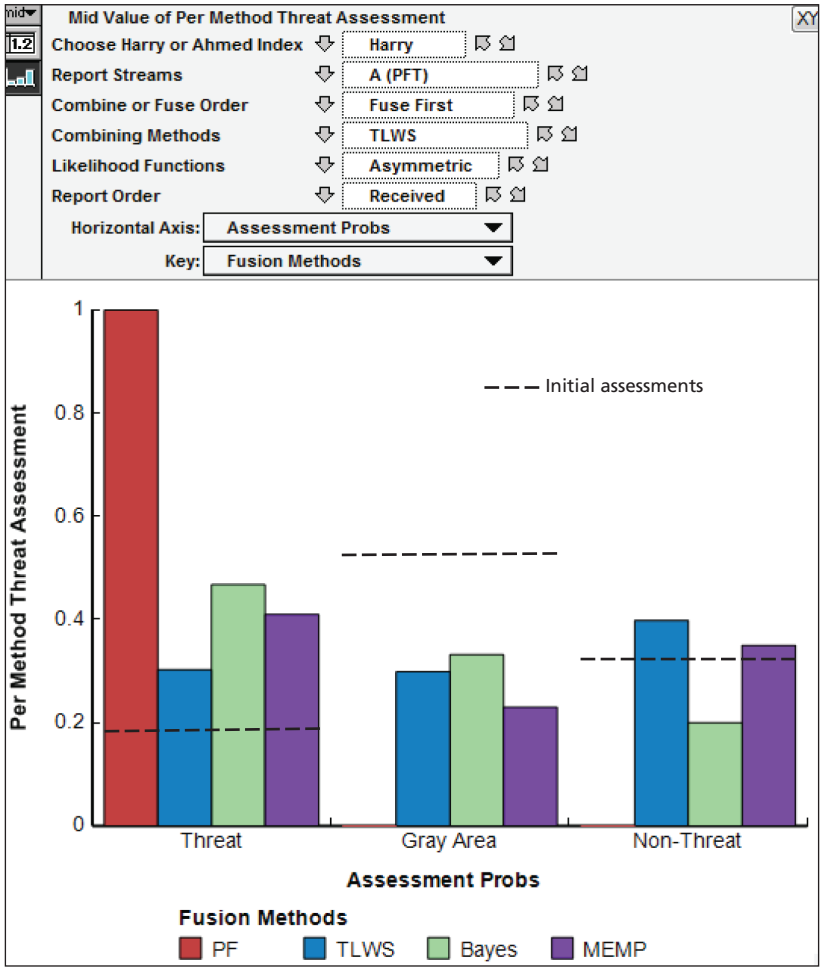
For this particular set of assumptions, the probability that Harry is a threat varies markedly, from about 0.3 to 1, depending on the fusion method used. The PF method is much more alarmist. The TLWS method is least alarmist, because it sees no threat unless *all* the factors of the PFT model exceed threshold values.

The last block of three bars shows related differences for the estimated probability that Harry is a nonthreat. Even if we put aside the PF result, which gives zero credence to Harry being a nonthreat, there remains a factor-of-two variation.

Ordinarily, analysts hope that calculations done in different ways will lead to the same results: The intent is just to check for errors. In contrast, we are comparing across methods specifically because they may give different results. When they do, it is necessary to look more deeply at the appropriateness of the methods and assumptions and to understand how sensitively results depend on data that are uncertain or even suspicious. Figure 6.2 tells us that such iterative analysis is needed for the Harry case.

This example also illustrates how an individual who did not initially appear to be much of a threat can be seen as a more obvious con-

Figure 6.1
Fusion Raises Likelihood That Subject Is a Threat



RAND *RR1200-6.1*

cern after fusion.* Thus, as anticipated, fusion may help detect threats. Later examples show how fusion may help to exonerate.

Returning now to the numerous variables held constant in Figure 6.1, we might ask what happens if the fusion were accomplished with the combine-first approach, rather than the fuse-and-combine approach. That is, each report estimates threat and the threat estimates are then fused across reports. The combine-first approach might be easier organizationally because it requires less careful comparison. Fusing first, however, has theoretical advantages, especially when individual reports have only fragmentary information about such factors as motivation and capability. Figure 6.2 compares results for the two sequences. The differences are modes in this case, but large in others.

Although summary displays such as Figure 6.2a and 6.2b are convenient for bottom-line results, analysts need more detail for some purposes. Figure 6.3 shows the probability density and cumulative probability distributions for the same case as in Figure 6.1. As is most clear in Figure 6.2b, except for the case of PF fusion, fusion assigns considerable probability mass to Harry being no threat at all. Other displays, of course, go into details of the underlying calculations.

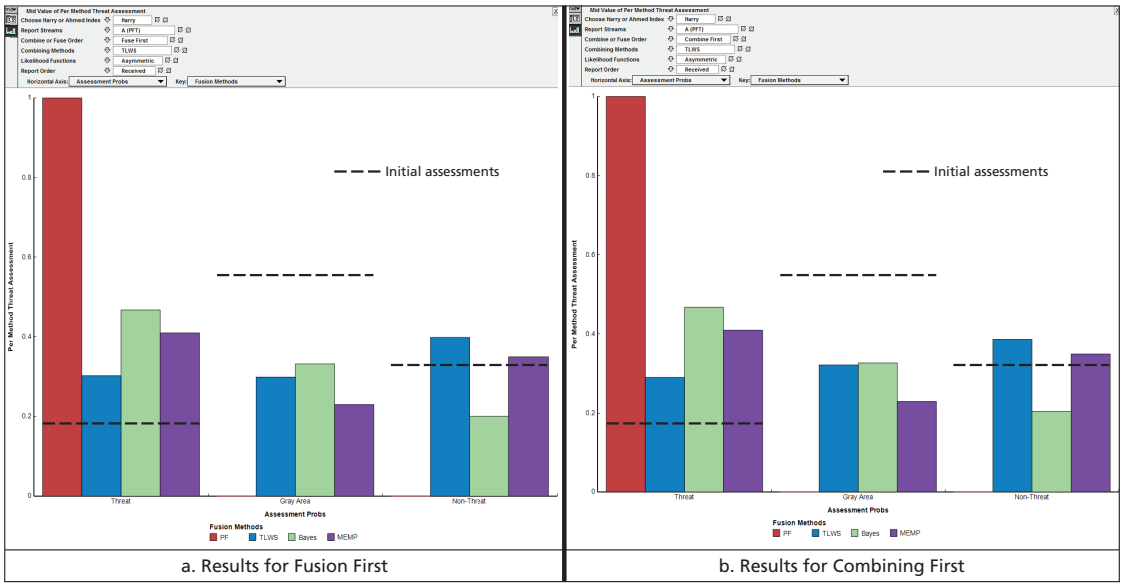
Initial Observations from Working with MEMP

Our MEMP fusion method is new (although building on much prior work, as discussed in Chapter Five), so we show only a few illustrative results with a different character than that of the other methods. Figure 6.4 is based on an MEMP analysis with particular assumptions about tuning parameters, as discussed in Chapter Five.† In this particular analysis, unlike most of those done in our short experimentation phase, MEMP was used to find a self-consistent entropy-maximizing

* The values after fusion reflect both having new information and the process of fusing reports, not just the latter. We can disentangle these because the effect of the new information is arguably just the simple averaging of the reports, which happens here to be the same as the TLWS result. Thus, for this chart, the value of fusion itself per se is the difference between the results for the other methods and TLWS. Harry is seen as between 1.3 and 3 times more likely to be a threat after fusing the information.

† The graphics are different because our MEMP work was accomplished in Microsoft Excel rather than Analytica.

Figure 6.2
Illustrative Comparison of Results by Method (Means Only)

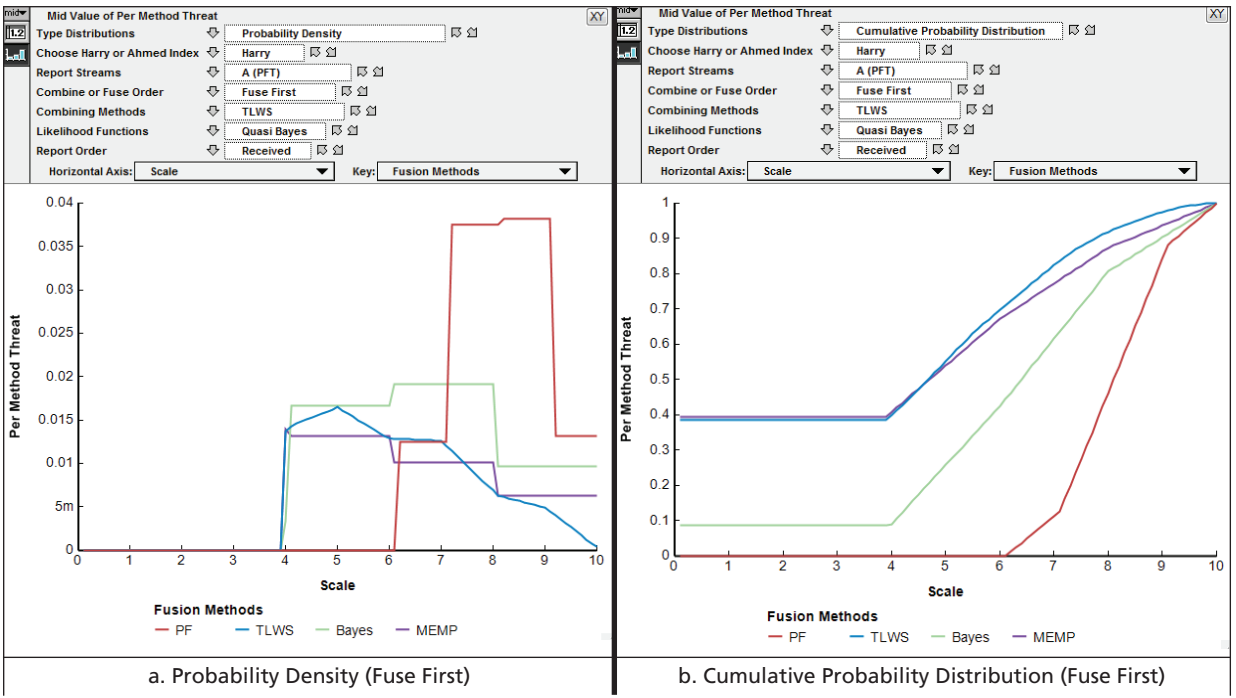


a. Results for Fusion First

b. Results for Combining First

RAND *RR1200-6.2*

Figure 6.3
Underlying Probability Distributions if Factor-Level Information Is Fused First



a. Probability Density (Fuse First)

b. Cumulative Probability Distribution (Fuse First)

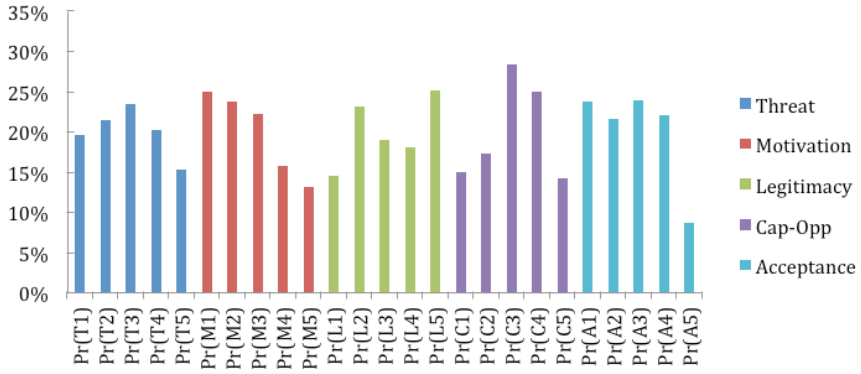
RAND *RR1200-6.3*

set of probability masses for the final threat estimate T and the factor values of the PFT model (see discussion toward the end of Chapter Five).

The first set of five bars on the left of Figure 6.4 show the probabilities of Harry being in the 1st, …,5th tiers of threat level T (i.e., of representing a very low, low, medium, high, or very high threat). The fused threat probabilities are all fairly close to 20 percent, except for the top risk level (5), which is at 15 percent. The fused estimates of the individual risk factors (motivation, etc.) vary more, with estimated probabilities from under 10 percent to close to 30 percent. Thus, the fusion process led to a more coherent—although highly uncertain—result.

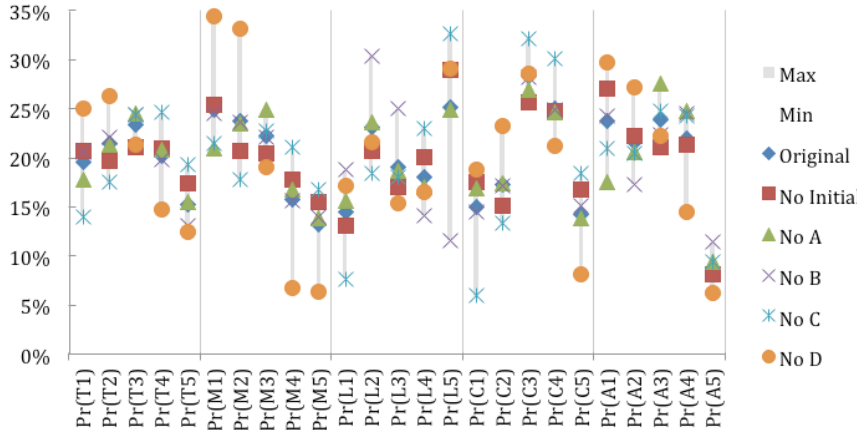
Our work on MEMP also generated interesting sensitivity analyses that we will later incorporate routinely with the others. Figure 6.5 shows the sensitivity of each probability estimate when one of the five reports (including the initial assessment) is dropped. For example, reading the leftmost column of symbols for $\Pr(T_1)$, the probability that Harry is a very low threat, we see that the probability increases to 25 percent if the report from Agent D is dropped. This is a straightforward way of recognizing quickly what reports are pivotal. This might tell us which report is most valuable or, on the negative

Figure 6.4
MEMP-Fused Estimates for Harry Assuming Equal a Priori Probabilities



RAND *RR1200-6.4*

Figure 6.5
Sensitivity to Dropping One of the Analysts' Reports



RAND *RR1200-6.5*

side, that conclusions depend sensitively on a single report of uncertain validity.

Other Explorations

Sensitivity to Usually Hidden Data

Figures 6.1–6.5 highlighted quite a number of variables in analysis, such as what fusion method to use. Other uncertainties, however, are usually hidden—a common problem in computer modeling. We illustrate this with one example drawing on the Ahmed vignettes of Chapter Three and, in particular, the role of the report by Ahmed's friend Abu. In addition to representing Abu's characterization of Ahmed as being definitely not a threat, the data from the vignettes include the analyst's assessment of the quality of the Abu report and an even more subjective assessment of the relative reliability of that to other reports of equal quality (see Chapter Two). Such data can be found in input tables, but such tables are deeply buried and often taken for granted. However, the architecture allows such inputs to be made much more visible rather easily.

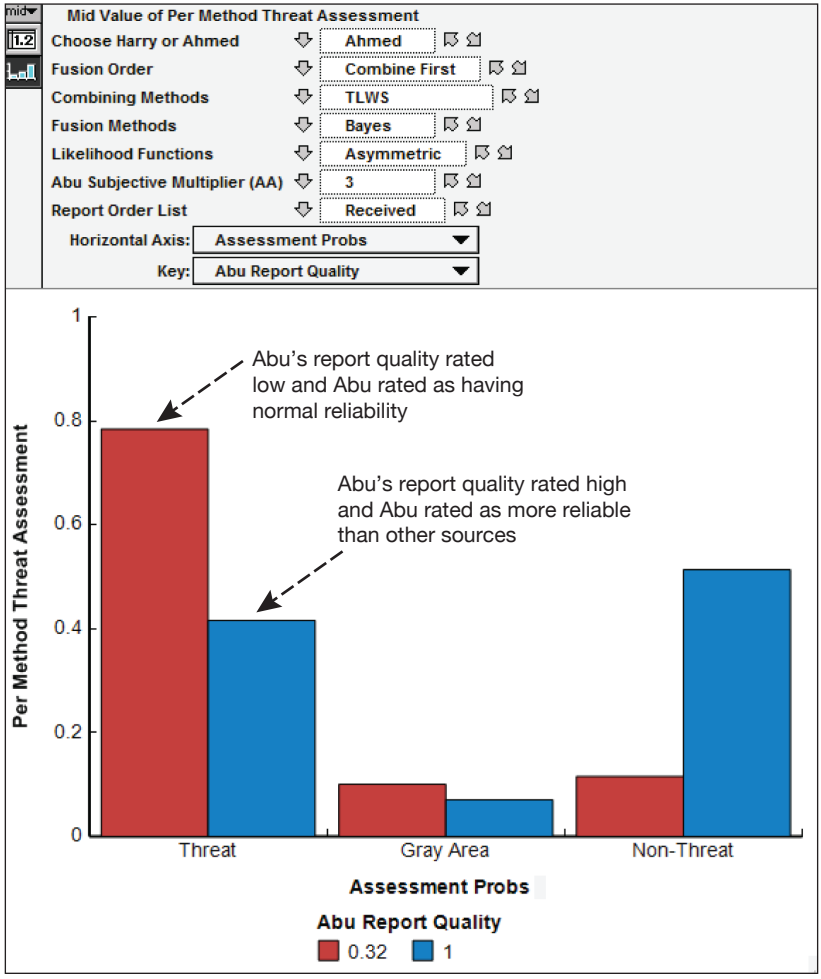
The quality of Abu's report was specified as 0.32 (low given that the datum is on a 0-to-1 scale) because of analyst concern that Abu was a friend of Ahmed. The analyst used the default assumption about reliability, which is that reliability should be assumed equal to quality. Suppose, however, that someone inquires on Ahmed's behalf, stimulating the review. The analysts would discover that changing the same variables highlighted in Figures 6.1–6.2 would have little effect on conclusions about Ahmed. They might then go back to original data and note that the report of Abu was at odds with the others. They might then elevate the visibility of the assumptions about quality and reliability. Figure 6.6 shows the consequences of doing so. Note the additional explicit variables for the quality of Abu's report and the subjective multiplier relating to reliability (additional "slicer bars").

The effect is that the likelihood of Ahmed being a threat goes from about 0.8 to about 0.4, a roughly 50 percent drop. These changes could come about because, in retrospect, the Abu case file was impressive: Abu has a clean history and holds security clearances; he went to a lot of trouble and arguably took risks in putting together strong testimony on Ahmed's behalf. Although a friend of Ahmed, he's not like a mother, obligated to protect Ahmed. His comments about Ahmed were clearly salient. Why, then, was his report given such a low rating? Why is it that Abu's report was considered no more reliable than that of a defector from terrorism (Mohammed, in the vignettes), who might have the attributes of true-believer converts or who might even be trying to impress the government so as to gain influence? The analysts might begin to worry that Mohammed was being duplicitous and Abu was being inappropriately discounted. If so, the consequences are strong, as shown in Figure 6.6.

As it happens, when we went back to the narrative in Chapter Three, the analyst's initial discounting of Abu's report still seemed to us reasonable (he, of course, was in reality one of us). That, however, was largely because of the negative information about Ahmed rather than any objective evidence against Abu's testimony. Did that constitute a bias?

The moral of the story is that fusion results can be sensitive to input assumptions that are buried in databases unless the effort is made

Figure 6.6
Sensitivity of Assessment of Ahmed to Assumed Credibility of Abu and His Report



RAND *RR1200-6.6*

to recognize such uncertainties. It is possible to elevate such sensitivities to the analyst's attention if the determination exists to do so. It may be that the questionable assumptions have little impact. In this case, however, they were important because the evidence as a whole is con-

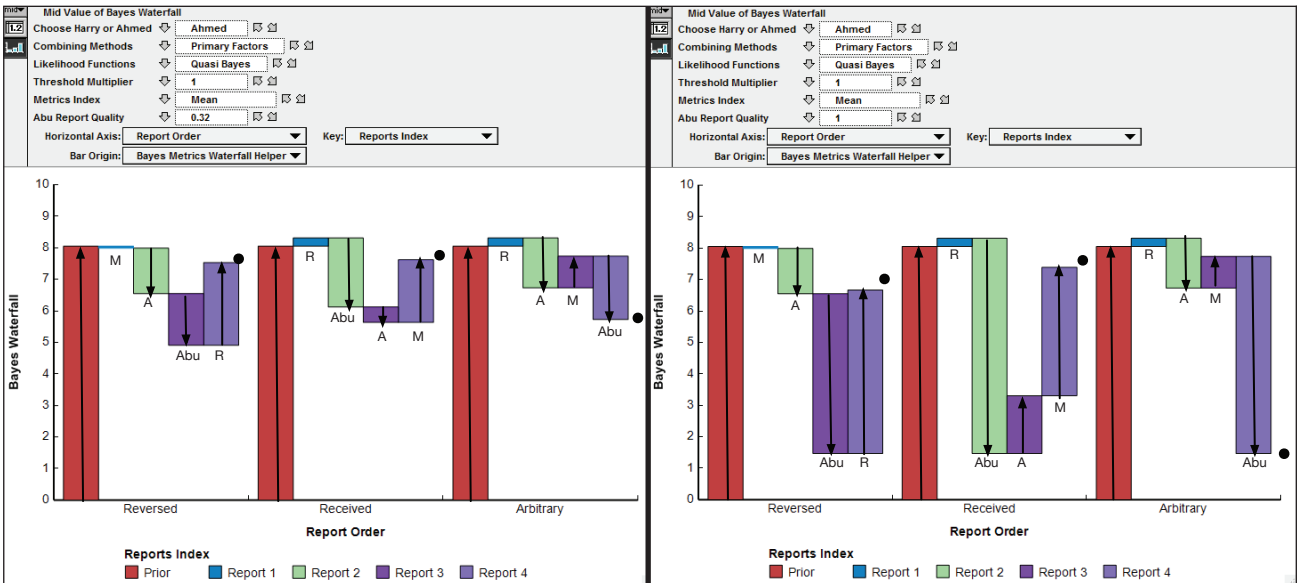
flictual, with no compelling evidence against Ahmed except that of a defector (albeit a convincing one). In this case, then, it would be important both to do further assessment of Mohammed and seek independent information. We hope that this example illustrates the importance of making uncertainty analysis easy and routine, and that it illustrates how uncertainty-sensitive fusion could in some case help to exonerate.

Questioning Details of the Analytic Process

As a next example of how the prototype platform provides flexibility to analysts wishing to view information from all angles, consider an issue often not even asked in technical work: Given a half-dozen reports that have come in over a period of time, how should they be processed? It might seem straightforward to process them in whatever order they arrive, although perhaps adjusting their reliabilities downward if they appeared stale or upward if they appeared to be very current. Figure 6.7 addresses the issue with the process of combining first and then fusing (not indicated in the figure), using TLWS and PF, respectively, and the quasi-Bayesian likelihood function. The mean is used when a single metric is needed to characterize a distribution.

Figure 6.7 is tricky to understand. We have annotated with arrows and legends. Looking at the leftmost group of five bars in pane (a), moving rightward allows us to see the threat estimate of the prior and to then move rightward for the other reports in the group. The arrows indicate whether a particular bar is raising or lowering the estimate relative to the previous one. The vertical axis is Ahmed's threat rating *T*. The horizontal dimension shows results along the stream of analysis if the reports are processed in (1) reverse order from that in which they are received, (2) the order in which they are received, or (3) in another arbitrary order that we chose for the sake of comparison. Order would not matter in a valid Bayesian analysis of empirical data, but in this application our quasi-Bayesian methods are using constant postulated likelihood functions that do not reflect cumulative knowledge and may be otherwise imperfect. As a result, order can matter. Further research is needed to fully characterize when order dependence occurs, and to consider how a more sophisticated Bayesian analysis could address these instances.

Figure 6.7
Changes in Threat Estimate with Processing of Reports in Different Order



NOTE: The abbreviations M, A, and R stand for the reports by Mohammed, Agent A, and Agent R, respectively. The first bar in each group is the same prior. The dark circle indicates the final result.

RAND *RR1200-6.7*

Looking at the first group, for processing in reverse order, the assessment of T is 8 at the time of the prior, with updated values of 8, 6.5 (downward arrow), 4.8 (downward arrow), and 7.6 (upward arrow) after the reports of Mohammed, Agent A, Abu, and Agent R, respectively. The final assessment of 7.6, then, places Ahmed in the high category of threat.

The second and third groups, corresponding to the as-received and a more arbitrary order, end up assessing Ahmed as 7.6 and 5.8, respectively. Taken together, the conclusion might be that Ahmed's threat score is in the medium-to-high range. Order of processing mattered, but not too dramatically. Pane (b) shows results with one change of assumption. Because order of processing is seen to have made a difference, the analyst has gone back to the underlying data to understand why (as imagined also in the previous section). He has noted that Abu's report, which insisted that Ahmed was no threat at all, had been given a low quality/reliability rating (as discussed in the previous section), an assumption that may not have been appropriate. The analyst now tests to see the consequence of increasing the assessed reliability of Abu's report to the same level as that of Mohammed. Now we observe *dramatic* effects depending on the order in which the reports are processed—a rather disquieting result that should give pause. In particular, if merely Abu's report is processed last (rightmost grouping), Ahmed is assessed to be no threat at all (a score of 1.5).

These problems reflect the approximation that makes our method "quasi-Bayesian" (see discussion in Chapter Four). The approximation favors using new information when it is more reliable than previous information, but downplaying it when it is less reliable. The heuristic for doing so creates order dependence in some cases. The consequence of the order dependence also depends on the assumed likelihood functions, which in principle should reflect accumulated knowledge from previous reports, rather than being constant. An important point of this example is to illustrate how exploratory analysis changing analytic assumptions and even "structural" assumptions, such as the order of reports, can reveal oddities telling the analyst to look more deeply into assumptions, and relationships among reports.

In this case, the quasi-Bayesian update was troublesome because the assumed reliabilities of the particular reports were such that the order dependence showed up. The problems were exacerbated because, for the particular run shown in Figure 6.7, the likelihood function used was unreasonably narrow given the knowledge available.* More generally, the quasi-Bayesian update is not obviously appropriate in this case because of the severe conflicts between reports: The core issue is deciding on the relative credibility of those reports, not precisely how to update from one to the other. This is easily recognized when one looks into the specific case and its data.

In future work, much more could be done to build in automated diagnostics to indicate when a given method is operating under circumstances where it is not valid.

Dealing with Bias and Related Correlations

Although we were able to do only a modest amount of research on the role of stories and other higher-level correlations in our first phase of research, the issue merits discussion, even if brief.

Stories

As discussed in Chapter One, mental stories play a significant role in fusion analysis. Table 6.1 illustrates briefly our current thinking about how to incorporate stories in the overall approach. Suppose that an analyst team has the same set of hard facts, but is having trouble interpreting them or estimating threat in the Harry vignettes. Two narratives have emerged about how to understand what is known about Harry's behavior, which includes occasional participation in meetings with a right-wing gang but also includes holding down a job and seeming normal to coworkers. The stories/narratives given in Table 6.1

* Chapter Four discusses how a generic distribution could be chosen at a particular point in a sequence of evaluations. Given accumulated knowledge, it might be possible to use a relatively "tight" likelihood function, whereas earlier that might be inappropriate. The order dependence discussed for Figure 6.7 can be seen as an artifact of having used an inappropriately tight and constant likelihood function.

Table 6.1
An Interface Model for Dealing with Stories

Story One	Story Two
He's frustrated and unhappy, but ultimately very moral and kind with a fine mother. He's got some bad friends and he likes excitement, but they're losers who never get things done. He'll outgrow them. Don't over-react to information.	He's very angry and blames society. He enjoys being with other such people and does risky things, either because he doesn't care or has no sense. He's easily led and his friends have weapons and access to targets. He's smart enough not to broadcast his intentions to casual acquaintances. We need to be very cautious.

reflect background information from indirect interviewing of relatives and acquaintances. Although both purport to be describing Harry, the contrast illustrates how different observers and even analysts might describe what they see—reflecting not just facts, but mental images (stories) about how to understand those facts.

If analysts were aware of the different stories and the need to avoid tilting inappropriately toward the more favored one, they might develop two sets of inputs (interpretations for information). Table 6.2 shows the kind of data table that might be produced. It assumes use of the PFT model, as discussed in earlier chapters. Table 6.2 is an example of a more general analytic technique called using *interface models*.*

Table 6.2
Possible Expression of Priors

Inputs	Triangular Distribution Parameters	
Factors, Thresholds	**Story One**	**Story Two**
Motivation	2,3,6	4,6,8
Legitimacy	2,4,6	4,7,8
Capability-Opportunity	0,3,7	5,7,8
Acceptability of costs	3,4,6	5,7,9

* Such interface models are typically necessary for exploratory analysis of policy problems. Many past examples exist (Davis, McEver, and Wilson, 2002; Davis, 2014a).

The significance of the different stories is shown in Figure 6.8, assuming just a single report for simplicity. The only fusion in this case is combining over the different factors using the PFT model with the TLWS method (in most of the report, we distinguish this type of fusion and call it "combining"). Pane (a) shows the resulting probability distribution for the individual's threat level. Pane (b) shows what might be thought of as a good aggregation: the probability that the individual's threat level falls in the range 6 to 10 (high or very high). We see that the results are markedly different for the two stories even though the underlying data was the same: The interpretations were different, and systematically so.

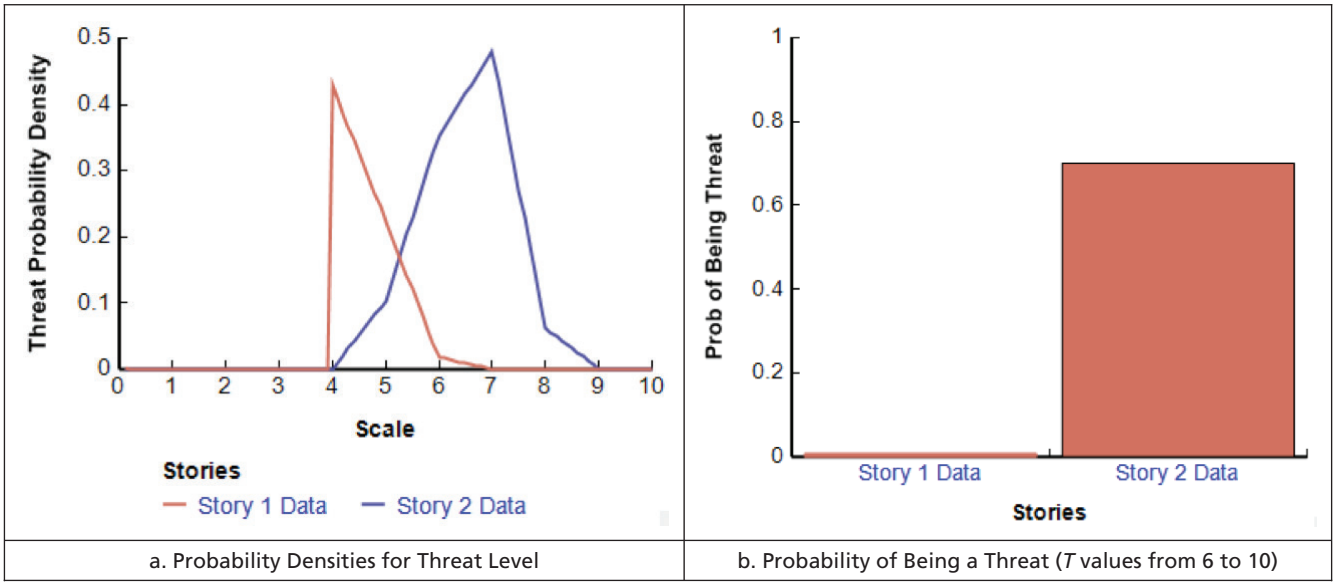
The moral of the story, of course, is that a fusion process should recognize where different stories are in play. The process should make it clear how results vary with story. This might seem trivial, but a notorious tendency in both law enforcement and intelligence is to focus attention on only the story that currently has most resonance. This is one of the many errors that intelligence analytic tradecraft seeks to avoid. Doing so is easier said than done, especially when choices have to be made about how to apply resources.

Other Sources of Correlation

Many other examples of correlated biases may come in at lower levels, in the day-to-day work of receiving, interpreting, and fusing reports. Some examples are, in generic terms:

1. Reports coming from different sources who observed the same events. If an event is reflected in three reports rather than one (perhaps with different names and details), that does not represent three times more information. In such cases, reports should be packaged or their weights reduced (see also discussion in Chapter Seven).
2. The sources of different reports may share beliefs about background facts and about the way the world works. These will affect subjective judgments and, if not accurate, will be introducing systematic bias.

Figure 6.8
Final, Fused Threat Assessment as Function of Story



a. Probability Densities for Threat Level

b. Probability of Being a Threat (*T* values from 6 to 10)

RAND *RR1200-6.8*

3. The sources of different reports may have personal-level biases that cause observed associations to be regarded as worrisome, not worrisome, or not credible. Some obvious examples might be that friends and family might testify more positively about an individual than his enemies or competitors.

One topic for future work would be to see whether a useful taxonomy of type correlations can be developed and whether it could be used to suggest exploratory analysis that systematically reinterprets report data to check against evidence of such bias. Analogous methods have been developed for other domains.

Next Steps on Stories and Correlations Across Reports

As we wrapped up this report, we were able to do some minimal experiments using interface models, alternative stories, and crude correlation functions to test the sensitivity of results to choice of stories and degree of correlation. Much more is needed. In time, interactive, iterative, analyst-machine operations should help discover stories and correlations, to construct and evaluate the consequences of new ones added ad hoc.

More generally, we have not yet begun to experiment with using the rich set of computer search methods that can be brought to bear at different levels of our analysis. Among those are the robust decision making (RDM) methods used heavily in other RAND policy research (Lempert et al., 2006). In those applications, reference is often made to "scenario discovery," which is having the computer find the sets of assumptions that generate outcomes that are, e.g., very good to very bad. An analyst can conduct such exploratory analysis manually by systematically navigating through the problem space to see patterns, as suggested earlier in the chapter. With more uncertain parameters, however, machine methods become increasingly valuable.

Summary Observations from Experiments

To recapitulate, we have demonstrated a number of flexibilities that we have built into our approach and prototype platform, all of which appear to be significant. These include, notably, dependence of fused assessments on

1. fusion method
2. causal model
3. relative order of fusing model-factor information and fusing threat estimates
4. subjective assessments of credibility and reliability
5. likelihood function
6. heuristic cut-off rules on quality
7. thresholds
8. whether and when to fuse across streams of analysis.

Which of these would be important in a given analysis is impossible to say. We would expect that some of these choices could be standardized after experience, reducing unnecessary degrees of freedom in fusion. That said, in looking at synthetic cases, we tend to see more rather than less reason to have such flexibility: The kind of capabilities that we are researching are not intended for use by tyro analysts applying mechanical methods, such as invoking a standard spreadsheet function to compute the statistics of least-squares data analysis.

The primary purpose of this chapter has been to demonstrate that the various fusion methods have been implemented and to illustrate how results of heterogeneous fusion can vary with the method used, and also with variations in process and relatively subtle assumptions. As noted throughout the report, we envisage fusion with respect to threat detection as being something that demands highly skilled analysts with good judgment and an appreciation for uncertainties—not only in data but also about modeling assumptions and how to go about processing information. It will take far more work to assimilate the implications of the methods and tools, and to better understand how they should best be used, but prospects for such investigation are very good.

CHAPTER SEVEN

Conclusions and Recommendations

As described in Chapter One, our objective was to design and illustrate, in prototype, an uncertainty-sensitive heterogeneous fusion process with the *potential* to (1) improve the likelihood of identifying actual terrorist threats, (2) reduce the number of erroneous identifications (false alarms), and (3) improve the ability to exonerate individuals previously under suspicion. With those broad objectives in mind, the more particular and modest objective in our research was to design, implement, and conduct initial tests with a prototype system that would do the following:

1. Address matters probabilistically rather than seeking point conclusions.
2. Take a mixed-method approach to information fusion.
3. Allow for parallel, competitive, and iterative streams of analysis.
4. Employ causal models rather than just statistical data.
5. Routinely use exploratory-analysis methods to deal with uncertainty affecting models, process, and assumptions.

For reasons discussed in Chapter One, such a system—*if feasible and practical*—should have the potential sought. Achieving the potential, of course, would depend on the quality of information, methods, and analysts—topics for subsequent research.

As discussed in Chapter Six, we were able to demonstrate the intended concepts and methods using our prototype platform and the synthetic data we developed. We were able to represent a broad range of complex probabilistic information that can be uncertain, fragmentary,

soft, contradictory, and even deliberately misleading. We developed and implemented first versions of fusion methods in several categories: algebraic, quasi-Bayesian, and entropy-maximizing. We used an illustrative causal mode with a variant (illustrating minimally a parallel stream of analysis). We demonstrated routinized uncertainty analysis in many dimensions.

The approach accomplished another objective, which was to improve the rigor of analysis by making explicit numerous assumptions and choices and by allowing repeatability, review, and variation.

The particular approach we took followed two paradigms that are worth noting here. They can be seen as reflecting what we saw as desirable principles for this particular application area with all its complexities:

- *A Mixed-Methods Approach with Analytic Self-Awareness.* The approach envisions routinely applying a number of different methods and assumption sets when attempting to fuse information. If the results agree, then the analysis is complete. If not, more work is needed. The fusion analyst can iterate with the benefit of the uncertainty-analysis tools: thinking more deeply about which methods and assumptions make the most and least sense for the particular case, reviewing the particular data shown to be especially important to the result, and sharpening an understanding of residual discrepancies and how they depend on discrete assumptions. This approach is in contrast to working with a single method and nominal assumptions, or that plus only modest sensitivity analysis.*
- *Human-Centric.* The approach recognizes the central and distributed role of uncertain judgments by sources, intelligence analysts, and fusion analysts throughout the process. This is in contrast

* An alternative would be to apply only the methods and assumptions most appropriate to a case. That, however, seemed impractical, requiring much human-intensive expert effort that would be wasted if results were overdetermined and requiring the ability to evaluate methods and assumptions up front without the benefit of initial experiments with the data. Future automated diagnostics could help make such judgments, but "playing with the data" to appreciate its complexities will continue to be important.

to an emphasis on data-driven automated analysis with machine learning and relatively minimal human interaction.
- *Ubiquitous Uncertainty Analysis.* Sensitivity analysis is a long and hallowed tradition, but in the domain of our study the uncertainties are "everywhere" and exploring the consequences of different assumptions and choices needs to be routine.

Our approach, then, is intended to supplement others, particularly the powerful data-driven approaches that are being pursued currently. Both are needed.

This said, our research to date has been basic in nature and should be seen as first steps in a process. We hope in future work to be able to do the following:

1. Experiment more with the prototype to better assimilate its lessons.
2. Sharpen and extend our methods (e.g., to include a Bayesian-network method).
3. Fill in theoretical gaps (e.g., understanding how best to use empirical base-rate data and the results of data-driven fusion exploiting data mining and machine learning).
4. Develop a more mature prototype platform that can be shared.
5. Use more comprehensive and representative data sets informed by real-world and exercise experience in the government.
6. Conduct experiments with analysts experienced in current-day high-level fusion operations in government.

If such work is successful, then it will be time for relevant agencies to move toward applied development. The ultimate value of the approach will then depend also on refining the quality of the models and methods and educating analysts in their use. Even then, of course, effectiveness in real-world applications will depend on the quality of the information available.

APPENDIX A

Defining Threat

Overview of Threat Concept

Suppose that an agency responsible for preventing acts of terrorism evaluates an individual who may commit or support a terrorist attack. The individual is characterized by threat T, a probability distribution that measures the importance that the agency sees in doing something about the individual, whether to investigate, put under surveillance, interdict, or arrest. The importance of action increases with the likelihood of the individual conducting some terrorist attack, the consequences of such an attack if successful, and the vulnerability to such an attack. The evaluation of T must also reflect the fact that the agency's assessment is looking for potential dangers that it has responsibilities to address. It cannot fail to note when someone might be a threat, even if he is probably not. That is, the assessment is not just about expected values (means) of some calculation.

The challenge is to turn this relatively simple intuitive concept into something tighter. This appendix sketches *conceptually* how that could be done.

Definitions and Conventions

Conceptually, threat T *could* be seen as the result of considering a vast array of possible attacks, the likelihoods of the individual attempting them, the likelihood of their success, and their consequences if successful. Threat could then be defined by a multidimensional integral over

all the uncertain variables with joint probability distributions. Given the impossible complexity of such calculations, we instead define T as in the model of Figure A.1, a highly aggregated construct that can be operationalized.* It considers attacks as falling into five different consequence classes with nominal scores of 1, 3, 5, 7, and 9. The variables are

T: threat posed by individual

T_i: the threat posed by individual with respect to attacks at consequence level i

W_i: the willingness (expressed as probability) to attempt level-i attacks

V_i: vulnerability of level-i targets (expressed as success probability for an attack)

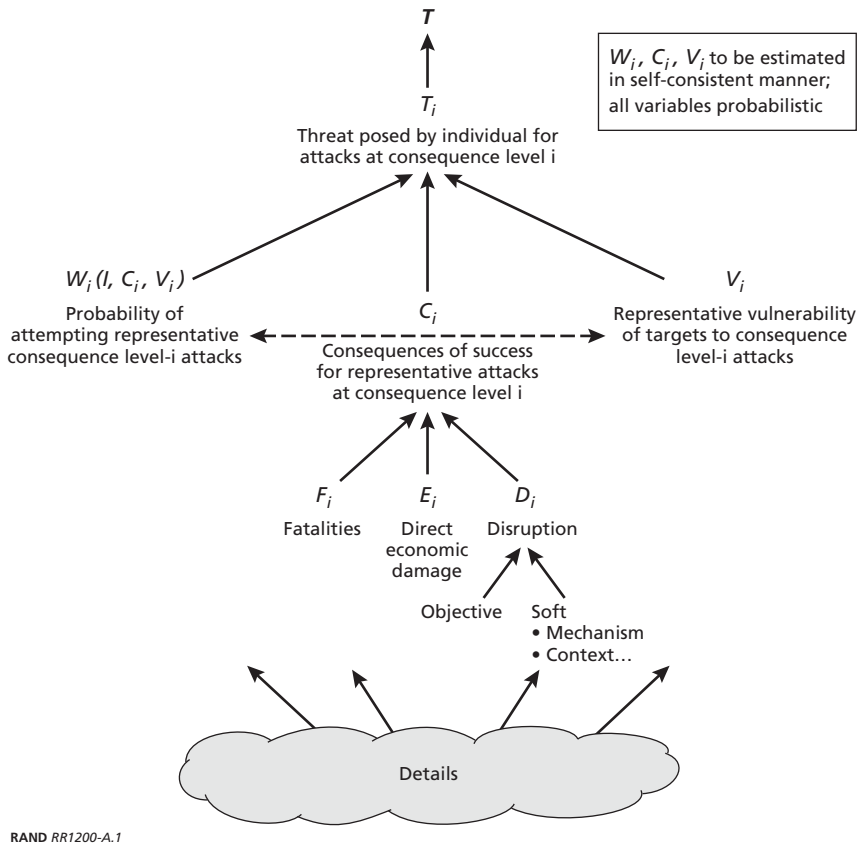
I: characteristics of the individual under study

C_i: consequences of successful level-i attacks (an aggregation of fatalities, direct economic damage, and disruption, with disruption effects of types called objective [easily measured] and soft).

As in Figure A.1, T is some aggregation of the T_i—threats posed at different consequence levels. These are approximated roughly as the product of the probability of a *representative* level-i attack, success probability, and consequences of success. This product form is ubiquitous in risk analysis, but is a *substantial* simplification. W_i, C_i, and V_i must be reasonable abstractions of lower-level detail. Their values must somehow be estimated in a self-consistent manner, because which attacks an individual considers depends on the value he sees in attempting the attacks, the vulnerability of targets, and related consequences. The approach amounts to assuming that a complex multidimensional integral can be approximated as such a product.

* Some risk-focused simulations assume low-level details and then assume doctrinal behaviors and optimizations along the way. The result may be unique and precise, but is fragile to buried assumptions (Morral et al., 2012).

Figure A.1
Conceptual Model of Threat



Specifications

The Individual. We assume that information on the individual is available from various sources and is sufficient to indicate, roughly, the kinds of interests, ambitions, and capabilities the individual would have for attacks.

Consequences. Turning to consequences, we can illustrate instructions on how to estimate values. Table A.1 assigns values to fatalities (*F*) and direct economic damage (*E*) on the importance scale from 0 to 10. For example, $1 million–$20 million of economic damage is rated

Table A.1
Assigning Importance Levels to the Elements of Consequences (Illustrative)

Type Consequences	Consequences (0–10)	Consequences (qualitative)
Fatalities		
0	1	Very low
1–10	3	Low
10–50	5	Medium
50–100	7	High
500+	9	Very high
Economic Damage		
$0–$1 million	1	Very low
$1 million–$20 million	3	Low
$20 million–$100 million	5	Medium
$100 million-$10 billion	7	High
$10 billion+	9	Very high

as a 3. This is deemed low from the viewpoint of an imagined agency focused on national counterterrorism.

The disruption variable (D) is more problematic. Even the relatively objective component is difficult to estimate, as demonstrated by the estimate range for attacks on the World Trade Center and Pentagon. The softer component is even more problematic, but important. The best that can be done is probably to provide a set of type events, give them standardized scores applying to typical examples, and expect the analyst to use them in the context of a particular evaluation.

Table A.2 lists a number of attacks and standardized disruption scores. The word *standardized* is important because, after some attack, the disruption that has actually occurred might be larger or smaller for innumerable reasons. The examples reflect our experience observing practice in law enforcement, intelligence, and counterterrorism. Establishing a firmer base for such scores would require a significant empirical effort.

Table A.2
Examples for Use in Characterizing Disruption Levels

Attack	Disruption (Qualitative)	Disruption (Importance Scale, 0 to 10)
Nuclear, biological, or radiological attack in city	Very high	9
Chemical attack in city	High	7
Bombing of public events	High	8
Bombing of critical infrastructure	High	7
Major assassination	High	7
Airliners attacking buildings (9/11)	High	7
Multiple armed assaults in populated area	Medium	5
Bombing iconic monuments	Medium	5
Common crime (even felonies)	Low or very low	1–3

NOTES: The scores measure potential disruption and assume sizable attack magnitudes. These may be regarded as modal values in probability distributions.

Although more complex algorithms are possible, overall consequences (C) could simply be calculated as

$$C = Max(F,E,D)$$
$$D = Max(D_{obj},D_{soft}).$$

Vulnerability. Vulnerability estimates should be subjective probabilities of successful attacks accounting for the terrorist's capabilities and values. Although a crude specification, this is more practical than may be evident. It can be straightforward to distinguish someone potentially likely to attack a government office building with a small ad hoc bomb from someone who might attempt a mass-destruction attack. Similarly, it can be straightforward to estimate the success

likelihood of the most plausible attacks based on his access to targets, explosives, etc.*

Willingness to Attack. The variable W_i must be estimated for the particular individual in question, accounting for considerations such as motivation, capability and opportunity (for the most plausible targets in each consequence level), and his understanding of both success probabilities and risks. W_i should be expressed as a probability of attempting an attack of level i.

Functional Form for Estimating Threat at a Given Consequence Level

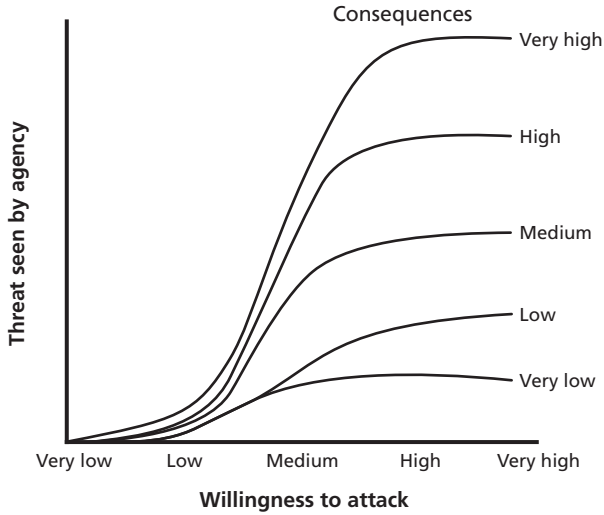
Figure A.2 specifies schematically the functional form for evaluation of T_i. It shows T_i as a function of willingness and consequences. The functional relationship sketched is nonlinear because an agency will see it as important to act if feasibility is greater than very low.

The next issue is the function determining T. It is assumed plausible that an individual might attempt attacks at different consequence levels. The aggregate concept of threat, T, should reflect the agency's desire to be especially cautious about individuals who may pose higher-consequence threats. Many functional forms could accomplish this, such as a linear weighted sum with more weight given to higher-consequence items.

Specifying the Probability Distributions. The variables of this conceptual model are probabilistic. We cannot realistically derive the distributions from microscopic variables, but it is possible to attach reasonable distributions. For example, because of uncertainties and random effects, C_i might be input as triangle (4,5,8) rather than as a point value of 5. In fusion problems related to detecting terrorists, the tails are important and it is better to approximate them than to ignore them.

* Some agencies expend considerable effort to characterize possible attacks and vulnerabilities. See, e.g., work by the U.S. Coast Guard supporting its Maritime Security Risk Assessment Model (MSRAM), available on the web.

Figure A.2
Threat as a Function of Aggregate Consequence and Vulnerability (Schematic)



RAND *RR1200-A.2*

A Highly Simplified Example

Suppose that an individual known to be highly motivated to conduct a terrorist attack but is regarded as conceivably capable of only two types of attack. Suppose that, on the 0-to-10 scale, the first attack would have potential consequences 5 and the second would have potential consequences 8 (e.g., an attack with a group that would seek to kill a number of civilians in a public gathering, using guns, or an attack with a group attempting to kill many more civilians in a public gathering using an improvised area weapon, such as in the Boston Marathon event). Whether based on potential fatalities or potential disruption (Tables A.1 and A.2), consequences would be in the 7 to 8 range. If the agency was doubtful about the individual's ability, assigning a success probability of perhaps 0.5, then Figure A.2 indicates that threat would be in the same 7 to 8 range. If probability distributions were used, perhaps T would be characterized as a triangular distribution (triangular, 3,7,8). That would indicate a nontrivial possibility that the

individual would never get around to attempting an attack or might readily fail, but—because of the potential consequences and the *potential* for success—most of the weight would be in the high-threat range.

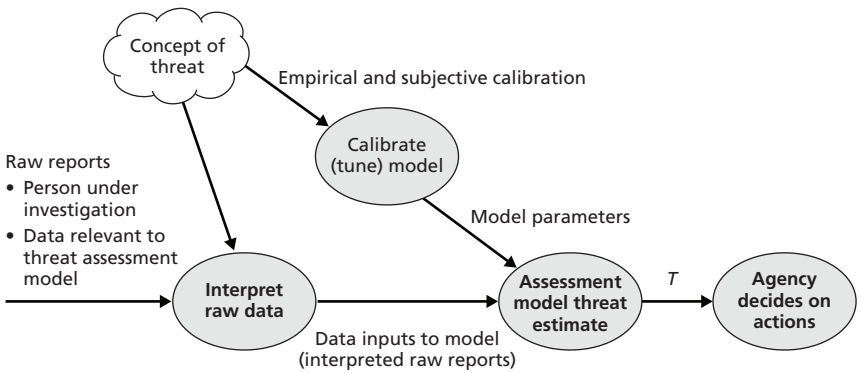
Implications for Threat-Assessment Models

The discussion above has been about the *concept* of threat. Given a particular threat-assessment model, such as the Propensity for Terrorism (PFT) model described in Appendix B, how do we use the abstract concept to inform PFT's structure, calibration, or use, and to assess whether it is sufficiently valid? Figure A.3 suggests the issues.

Reading left to right in Figure A.3, information comes in, including past reports and new raw reports that must be interpreted for model-based analysis based on a shared concept of threat (top left cloud). Once the model is calibrated to the same concept of threat, inputs from reports allow the model to generate an estimate of threat T appropriate to the agency in question. The agency uses the threat estimate T (along with other contextual information) to make decisions about actions.

Treating the concept of threat described in earlier sections as the base, how might the validity of a model such as PFT be assessed?

Figure A.3
Relationship Between Concept of Threat and an Assessment Model



RAND *RR1200-A.3*

How might it be tuned to generate threat estimates consistent with intent? Actually doing such tuning and validation would go far beyond the scope our current research, but some observations are possible, illustrating them for the PFT model.

- If we used versions of PFT that applied simple linear weighted sums, the results would be inconsistent with the conceptual threat model: Someone with very high motivation, no doubts about legitimacy, and willingness to accept the costs and risks of terrorism would be assessed as a high threat even if he had no capability. The *thresholded* linear weighted sum (TLWS) method is thus more appropriate. Further, the threshold should probably be medium or perhaps low, but not very low.
- If, in evaluating the *CO* variable, we assigned a very high value because the individual was quite capable of very low-level terrorism, then in some cases we would generate a threat estimate far higher than intended by the conceptual threat model. It follows that we should calibrate the *inputs* to PFT to be consistent with the conceptual model. In the instance of an individual potentially willing and able to conduct only low-level terrorism, then, PFT's threat assessment should never exceed medium (4–6).*
- Because the conceptual threat model sees importance grow faster than the probability of success in an attack, and since there is no structural parallel within the TLWS version of PFT, it follows that those estimating *CO* should arguably err on the high side. For example, the threshold used for *CO* should be no greater than medium.

Even this brief discussion should demonstrate that calibrating a threat assessment model and the protocols appears feasible but nontrivial. Further, having an underlying conceptual model, as described in this appendix, is useful.

* As we learned in our experiments, calibration at the time of data entry can be troublesome because the sources for raw reports on an individual will tend to see high or very high threat levels even if, objectively, the attacks of which the individual might be capable fall short of, say, 9/11, the Boston Marathon bombing, or attacks with weapons of mass destruction. This suggests the need for a final check by the fusion analyst.

APPENDIX B

A Factor-Tree Model for Propensity for Terrorism (PFT)

Background

Earlier public-source RAND research included a voluminous critical review of the social science of counterterrorism (Davis and Cragin, 2009) and a subsequent in-depth look at the social science of public support for terrorism (Davis, Larson, et al., 2012). A novel feature of this work methodologically was its introduction of "factor-tree models," deceptively simple graphs showing the primary causal factors influencing the variable of interest (e.g., the likelihood of an individual becoming a terrorist or the likelihood of the public supporting a terrorist organization and its actions). Such factor trees are *qualitative multiresolution causal* models, rather than statistical models. The 2012 volume refined the factor tree for public support of terrorism and then exposed it to empirical testing with four new case studies. The factor tree held up quite well, which should not be surprising: The factors were identified from a large body of high-quality qualitative social science. Experts, such as the authors of the underlying social-science literature, are quite good at identifying the factors at work even if, as is well known, they are often not especially good at predicting the consequences—i.e., in describing combined effects of the factors.

A later study constructed a first-of-its-kind computational model based on a factor tree (Davis and O'Mahony, 2013). The authors emphasized that the result, the Public Support of Terrorism Model (PSOT), should not be used for prediction: Its factors have uncertain values, and there are significant questions about exactly how they com-

bine. Thus, PSOT is an uncertainty-sensitive model that allows exploratory analysis to better understand relationships, to see different ways by which results could arise, and to recognize desirable and ominous patterns of factor values. That can be valuable for decision-aiding even when uncertainties are large. The causal-model philosophy used has an interesting linkage to descriptions of alternative pathways by the late Alexander George in pioneering work on structured case studies in social science (George and Bennett, 2005). To modelers, factor trees can be seen as a static simplification of influence diagrams. The Davis-O'Mahony report includes an appendix that is a primer on factor trees (originally published as Davis, 2011).

Such factor-tree models and the computational model PSOT may seem odd to an economist, mathematician, or decision analyst familiar with classic rational-actor formalism. Such models attempt to represent phenomena more insightfully than is possible with rational-actor methods. For example, they recognize that people may be caught up with a movement or exciting activity and do very unwise things. They may act from idealism, religious faith, honor, or, bluntly, excitement and bloody-mindedness.* The PSOT and PFT models treat rational-actor behavior as a special case.

The PFT Model

For our research on uncertainty-sensitive heterogeneous fusion, we developed the Propensity for Terrorism (PFT) model. We did so as a spin-off of the earlier work cited above, drawing on the same base of social science. Although not separately documented or validated, PFT should have roughly the same validity as the earlier PSOT model. It was more than sufficient for our research purposes in that it illustrates

* Shortcomings of the rational-actor model are discussed elsewhere (Morgan, 2003; Davis, 2014b; National Research Council, 2014). That humans do not behave according to the model has been exhaustively demonstrated by Nobel Prize winners (Simon, 1978; Kahneman, 2002). Further, decisionmakers often do better with more naturalistic decision-making approaches or a hybrid (Klein, Moon, and Hoffman, 2006a, 2006b; Gigerenzer and Selten, 2002; Davis, Kulick, and Egner, 2005).

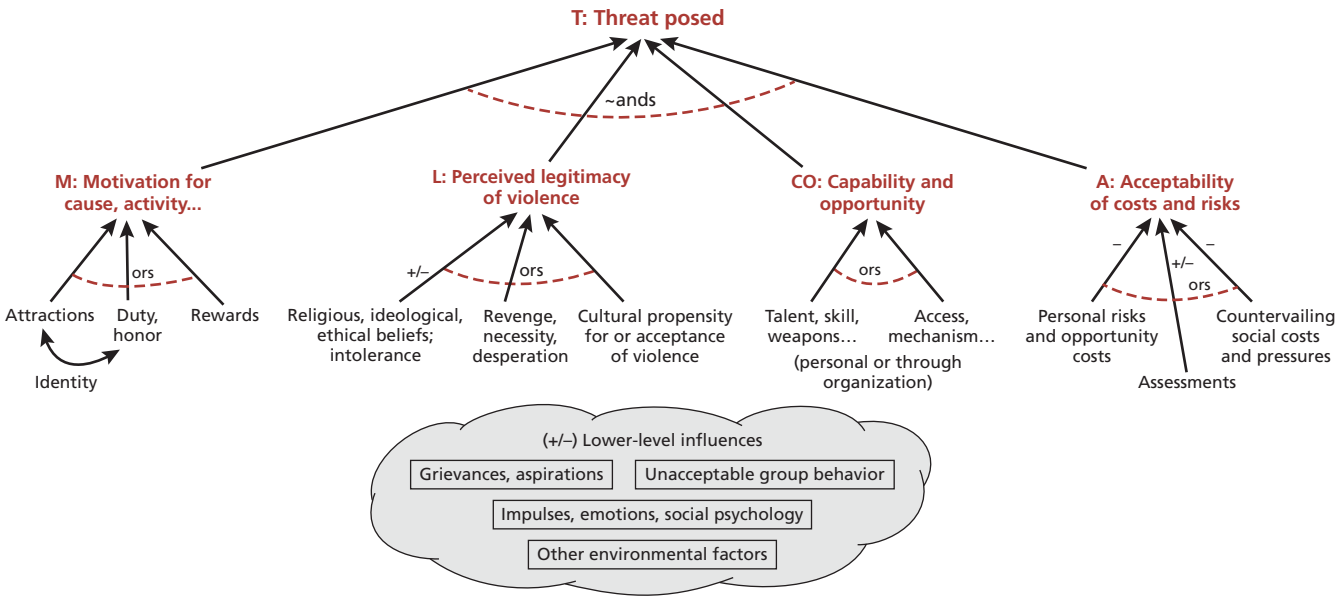
how a causal model rooted in social science can be used as an important component of fusion.

Figure B.1 describes the PFT factor tree. As shown here, it has less detail than that of the PSOT tree in the references above. The salient feature is that it describes the propensity for terrorism, which is mapped into the threat variable *T* discussed throughout the report, as a function of four independent factors: *M*, *L*, *CO*, and *A*:

- *Motivation (M)* measures the degree to which the individual is motivated by the relevant cause or activity. The value of *M* is the result of one or many lower-level influences that may include religious zeal, hatred of a suppressive government . . . or a love of dangerous and exciting action with comrades.
- *Legitimacy (L)* measures the degree to which the individual sees terrorism violence as justified, above and beyond the effects of motivation. *L* may be low even if the individual is strongly motivated for the cause or action because, for example, he may see terrorist violence as immoral or as severely counterproductive to the cause.
- *Capability-Opportunity (CO)* measures the degree to which the individual has the capability and opportunity to successfully execute relevant terrorist violence.
- *Acceptability of Costs (A)* measures the degree to which the individual is willing, perhaps for the sake of the cause or activity, which is a "positive," to accept the associated risks to his own life and that of families and friends, opportunity costs, and criticism by those he values.

The factor tree indicates with the approximate "and" condition at the top that for the individual to be a significant threat, *all* of these factors must be present. That is, *T* will be small unless the individual has motivation, perceives legitimacy, has the capability and opportunity, and is willing to pay the price. The approximation sign "~" is important (elsewhere in the report, regrettably, we use the same symbol to mean "not"). Social science is complex, and exceptions exist. Further, some of the factor values could change suddenly. Thus, the model's baseline

Figure B.1
Factor Tree for Threat (*T*) Posed by an Individual or Small Group Under Scrutiny



NOTES: 1. "ands and ors" apply strictly to binary case.
2. (+), –, +/–: influence is positive, negative, or a mix.
3. lower-level influences may affect multiple nodes.

RAND *RR1200-B.1*

structure is stringent, but uncertainty analysis is crucial. This should address uncertainties about factor values, "thresholds" (e.g., how much of a factor is necessary), and combining rules. All of that is enabled in the PFT model for information fusion described in the main text.

Our description here is conceptual. Details are described in the main text, Davis and O'Mahony (2013), and unpublished documentation. Although *M*, *L*, *CO*, and *A* are intended to be independent, reported information on them may incorporate unintended statistical correlations, as discussed in Chapter Three.

APPENDIX C

Extending Elementary Bayes Updating

This appendix considers an example of the elementary Bayes formula that we found useful in adapting Bayesian methods to our problem domain. As a variant of the classic coin-toss problem:

> Assume that you are evaluating the fairness of a coin being used at a party. You assume that the coin may be heads-only, fair, or tails-only, corresponding to a heads probability of 1, ½, or 0, respectively. At the outset, you assume that all three possibilities are equally likely. You now observe a trial with three flips of a coin with the outcome of three heads (HHH). What should your revised estimate be of the three possibilities?

Bayes' rule for the problem can be expressed as follows. The hypothesis set is

$$H: \{H_1, H_2, H_3\} \; ; \; H_1: \text{Heads-only}; \; H_2: \text{Fair}; \; H_3: \text{Tails-only}.$$

The appropriate variant of Bayes' rule is now an array equation that can be more easily understood by considering its component-level expression. For the ith component,

$$\Pr(H_i \mid E) = \frac{\Pr(E \mid H_i)\Pr(H_i)}{\sum_{k=1}^{3} \Pr(E \mid H_k)\Pr(H_k)} . \qquad \text{(C.1)}$$

This is an array equation, because the prior is a linear array in H (Table C.1) and $\Pr(E|H_i)$ is the appropriate column of Table C.2. The rows of Table C.2 show all possible outcomes of an experiment in which the same coin is flipped three times. The probabilities of those crisp outcomes depend on whether and how the coin is biased. For example, the outcome of three heads will occur: always, 1/8 of the time, or never, depending on coin bias.

Substituting into Equation C.1 the numbers from Table C.1 and C.2, the equation for the posterior based on an observation of HHH, for example, becomes

Table C.1
The Prior Distribution Expressed as a Linear Array

H1	H2	H3
1/3	1/3	1/3

NOTE: Cell values in the lower row are the prior probabilities of the hypotheses in the top row.

Table C.2
Likelihoods for Simple Coin-Flip Problem

Crisp Outcome E	Hypotheses H		
	H1 (Head Bias)	H2 (Fair)	H3 (Tail Bias)
HHH	1	1/8	0
HHT	0	1/8	0
HTH	0	1/8	0
THH	0	1/8	0
HTT	0	1/8	0
THT	0	1/8	0
TTH	0	1/8	0
TTT	0	1/8	1

NOTE: Cell values are probabilities of observing the row's outcome given the column's hypothesis.

$$\Pr(H_1 \mid E) = \frac{(1)(1/3)}{(1)(1/3)+(1/8)(1/3)+(0)(1/3)} = \frac{8}{9}. \quad \text{(C.2)}$$

The "(1)" in the numerator is the probability of the outcome HHH if the H_1 hypothesis (the heads-only case) is true. The "(1/3)" is the prior's estimate of the likelihood of H_1. The denominator adds the probabilities across hypotheses for obtaining the HHH result. It could be that the coin is fair (the middle term) or tails-only (the last term). It follows that if we observe the evidence of HHH in a trial of three coin flips (the first row of results in Table C.2), the Bayesian update is that the probability ascribed to H_1 is 8/9. The other posterior probabilities would be 1/9 and 0 (H_3 [tails-only] is disproved). This updated probability would apply only if the new evidence happened to be HHH. A different updated probability would be calculated for each of the other outcome possibilities (rows in Table C.2).

The initial example assumed a single trial with a crisp outcome—i.e., the sequence head, head, head for three flips of the coin. More typically, the evidence would be for a larger number of flips and would be expressed as the fraction of times that the result was a head. Now let us assume 100 coin flips, again of the same coin. Assume also that the hypotheses were more subtle, distinguishing among coins that are all "nearly" fair, with head probabilities of 0.4, 0.5, or 0.6. Such hypotheses would be useful in trying to infer the bias of a coin in a gambling situation in which the coin flipper might be using a slightly biased coin so that results would seem just a matter of luck. We would then have H: $\{H_1, H_2, H_3\}$, where the hypotheses are for head probabilities of 0.4, 0.5, and 0.6.

If the prior again assumed equal probabilities for the three hypotheses, then Table C.2 would apply with the new hypotheses. Instead of an extremely long version of Table C.2, we would use the binomial distribution to compute likelihoods of equivalent outcomes (e.g., HHT is equivalent to THH). A trial with 100 flips could be described by the fraction of heads observed, and the likelihoods would be as in Table C.3 with an admittedly silly level of precision.

Table C.3
Likelihoods

Number of Heads in 100 Flips	Tail Bias: Head Probability: 0.4	Fair Coin: Head Probability: 0.5	Head Bias: Head Probability: 0.6
45	0.047811	.048474	0.000829

NOTE: Values in the 2nd, 3rd, and 4th columns are binomial (100, 45), where "binomial" is the binomial distribution, which takes two parameters: the number of coin flips and the number of positive results (heads rather than tails).

Updating and rounding, we find that

$$\begin{aligned}
\Pr(H_1 \mid E) &= \frac{(0.048)(1/3)}{(0.048)(1/3)+(0.048)(1/3)+0} = 1/2 \\
\Pr(H_2 \mid E) &= \frac{(0.048)(1/3)}{(0.048)(1/3)+(0.048)(1/3)+0} = 1/2 \qquad \text{(C.3)} \\
\Pr(H_3 \mid E) &= \frac{(0)(1/3)}{(0.048)(1/3)+(0.048)(1/3)+0} = 0
\end{aligned}$$

In this case, the evidence allows us essentially to rule out the tails-biased coin but it cannot distinguish much between the likelihoods of the coin being modestly heads-biased or fair.

This example is about as demanding as seen in elementary accounts. Let us now step back and ponder, however. Even the "simple" coin-flip problem has real problems:

1. Perhaps a coin can land vertically by bouncing to the nearest wall and remaining upright. The actual results might have been 45 percent heads, 35 percent tails, and 20 percent on edge. We would then need to adjust the data to be 80 flips in the trial with 45 heads and 35 tails. The results would then indicate head-bias as most likely (0.4 probability).
2. The coin flipper might be a con artist switching among coins. If so, the mathematics in Equation C.3 is wrong, because the likelihoods are assumed constant.

3. The evidence might be imperfect because the umpire responsible for calling the results sometimes garbles his words to call a head a tail, or vice versa.
4. The recorder might be in the pocket of a gambling group and might tilt the results.
5. The whole calculation depends on knowing the binomial distribution. What if we didn't know the source distribution (as in [2], with its covert coin-switching)?

Our concern is obviously not with coin flipping, but these easy-to-understand examples illustrate complications that also loom large in the threat-assessment problem.

Abbreviations

LWS	linear weighted sums
MEMP	maximum entropy/minimum penalty
NIA	National Intelligence Agency
PF	primary factors
PFT	Propensity for Terrorism Model
PSOT	Public Support of Terrorism Model
TLWS	thresholded linear weighted sums

Glossary

Term	Meaning in This Report
Analytic architecture	Design for structured process of fusing diverse classes of information in multiple ways
Cases	Akin to a scenario, a case is the story over time of the assessment of the threat posed by a particular individual. A case consists of vignettes, events, and analysis as slices in time. In our usage, a case also means the set of all relevant input data.
Combining	Evaluating a function from its contributing factors (its independent variables)
Compound inputs	Inputs that are expressed with logical operators, such as "or," "and," or "not"
Confidence	A measure of an estimate's reliability as when threat T is subjectively assessed to be between 8 and 10 with 80% likelihood
Consequences	A multi-attribute measure of the negative results of an attack. The aggregate-level attributes used are fatalities, direct economic damage, and disruption.
Convergence	A decided narrowing of uncertainty around a most likely value as a case develops

Term	Meaning in This Report
Deterministic variables	Variables with fixed values in a given calculation.
Elicitation	Drawing information from experts to estimate variables of interest
Entropy	A measure of uncertainty regarding the state of a system; it is implied by a probability distribution for the states of that system. Alt.: a measure of the multiplicity of states consistent with knowledge.
Exploratory analysis	Analysis that systematically considers the combined effects of many or all relevant uncertainties simultaneously. Its distinguishing feature is that its objective is to understand the problem, formulate hypotheses, and explore impact of varying assumptions.
Factor strength	The degree to which a factor value is threatening, whether continuous or discretely represented
Factor tree	A graphical depiction of a model showing the dependent variable as influenced by causal factors (variables), which are influenced by lower-level causal factors
Factors	Highlighted independent variables
Fusion	Putting together, as in creating a threat estimate by combining different information
Fuzzy variables	Generalizations of ordinary variables with imprecise values

Term	Meaning in This Report
Heterogeneous fusion	Fusion across types of information, such as qualitative and quantitative, direct and indirect, objective and subjective, crisp or fuzzy, honest and deceptive
Independent variables	Variables that stand alone and are not changed by other variables.
Inputs	The independent variables of a model, which can be changed individually without changing the others, and which do not change if the other independent variables are changed
Interface model	A mechanism for mapping high-level questions into variables of a more detailed model or a model structured with different questions in mind.
Likelihoods	The conditional probability distribution of the evidence given the hypothesis, treated as a function of the hypothesis.
Posterior	The updated probability of a hypothesis after evidence has been received
Prior	The probability of a hypothesis before processing information at hand
Probabilistic dependence or correlation	Relationship between two variables such that the probability distribution of one depends on the value of the other
Probabilistic variables	Variables characterized by probability distributions
Quasi-Bayesian analysis	Analysis that uses Bayesian machinery but with simplified and approximate versions of some critical elements.

Term	Meaning in This Report
Random variables	A variable whose possible values are outcomes of a random process
Reports	Packages of interpreted data
Risk	A measure of adverse possibilities that considers threat, vulnerability of targets, and consequences
Soft	Difficult to measure, as with many human attributes and effects thereof (overlaps with qualitative, subjective, fuzzy, and ambiguous)
Stochastic process	A collection of random variables, representing the evolution of some system of random values over time.
Stories	Mental models that are often used to make sense of data, including "connecting the dots"
Stream of analysis	A continuing analysis over a series of reports by a particular team
Threat	The degree to which a subject is regarded as a potential danger to society
Threshold	The value of a variable beyond which the variable's effects on the subject of interest rises rapidly from zero
Uncertainty	Shortfalls in knowledge about the correct model or the values of variables in a given model
Vignette	See "Case"

Bibliography

American Civil Liberties Union (2015), "Terror Watch List Counter: A Million Plus." As of November 16, 2015:
https://www.aclu.org/terror-watch-list-counter-million-plus

Armstrong, J. Scott (2002), "Combining Forecasts," in J. Scott Clements, ed., *Principles of Forecasting: A Handbook for Researchers and Practitioners*, pp. 417–439.

Ash, Robert B. (1990), *Information Theory*, New York: Dover (by arrangement with Oxford University Press).

Atran, Scott (2010), *Talking to the Enemy: Faith, Brotherhood, and the (UN) Making of Terrorists*, New York: Harper Collins Publishers.

Banavar, Jayanth, Amos Maritan, and Igor Voklov (2010, January 22), "Applications of the Principle of Maximum Entropy: from Physics to Ecology," *Journal of Physics: Condensed Matter*, Vol. 22, No. 6.

Bergen, Peter (2013, December 30), "Would NSA Surveillance Have Stopped 9/11 Plot?" *CNN*. As of November 16, 2015:
http://www.cnn.com/2013/12/30/opinion/bergen-nsa-surveillance-september-11/

Bishop, Christopher M. (2007), *Pattern Recognition and Machine Learning*, Cambridge, UK: Springer.

Boström, Henrik, Sten F. Andler, Marchus Brohede, Ronnie Johansson, Alexander Karlsson, Joeri van Laere, Lars Niklasson, Maria Nilsson, Anne Persson, and Tom Ziemke (2007), *On the Definition of Information Fusion as a Field of Research,* Skövde, Sweden: University of Skövde, Informatics Research Centre.

Chiang, Yun-Wei, Peter P. Borbat, and Jack H. Freed (2005, December), "Maximum Entropy: A Complement to Tikhonov Regularization for Determination of Pair Distance Distributions by Pulsed ESR," *Journal of Magnetic Resonance*, Vol. 177.2, pp. 184–196.

Clemen, Robert T. (1989), "Combining Forecasts: A Review and Annotated Bibliography," *International Journal of Forecasting*, Vol. 5, pp. 559–583.

Clemen, Robert T., and Robert L. Winkler (2007), "Aggregation of Expert Probability Judgments," in Ward Edwards, Ralph F. Miles, and Detlof von Winterfeldt, eds., *Advances in Decision Analysis: from Foundations to Applications*, Cambridge, UK: Cambridge University Press, pp. 154–176.

Das, Balaram (1999), *Representing Uncertainties Using Bayesian Networks,* DSTO-TR-0918, Salisbury South Australia: Defense Science & Technology Organisation (DSTO).

Davis, Paul K. (2003), "Synthetic Cognitive Modeling of Adversaries for Effects-Based Planning," in Jonathan Kulick and Paul K. Davis, eds., *Modeling Adversaries and Related Cognitive Biases*, Santa Monica, Calif.: RAND Corporation, RP-1084. As of November 16, 2015:
http://www.rand.org/pubs/reprints/RP1084.html

——— (2011), "Primer for Building Factor Trees to Represent Social-Science Knowledge," in S. Jain, R. R. Creasey, K. P. Himmelspach, K. P. White, and M. Fu, eds., *Proceedings of the 2011 Winter Simulation Conference*.

——— (2012), *Lessons from RAND's Work on Planning Under Uncertainty for National Security*, Santa Monica, Calif.: RAND Corporation, TR-1249-OSD. As of November 16, 2013:
http://www.rand.org/pubs/technical_reports/TR1249.html

——— (2014a), *Analysis to Inform Defense Planning Despite Austerity*, Santa Monica, Calif.: RAND Corporation, RR-482-OSD. As of November 13, 2015:
http://www.rand.org/pubs/research_reports/RR482.html

——— (2014b), *Toward Theory for Dissuasion (or Deterrence) by Denial: Using Simple Cognitive Models of the Adversary to Inform Strategy*, Santa Monica, Calif.: RAND Corporation, WR-1027. As of November 19, 2015:
http://www.rand.org/pubs/working_papers/WR1027.html

Davis, Paul K., and Kim Cragin (2009), *Social Science for Counterterrorism: Putting the Pieces Together*, Santa Monica, Calif.: RAND Corporation, MG-849-OSD. As of November 17, 2015:
http://www.rand.org/pubs/monographs/MG849.html

Davis, Paul K., and Paul Dreyer (2009), *RAND's Portfolio Analysis Tool (PAT): Theory, Methods, and Reference Manual*, Santa Monica, Calif.: RAND Corporation, TR-756-OSD. As of November 13, 2015:
http://www.rand.org/pubs/technical_reports/TR756.html

Davis, Paul K., Jonathan Kulick, and Michael Egner (2005), *Implications of Modern Descision Science for Military Decision-Support Systems*, Santa Monica, Calif.: RAND Corporation, MG-360-AF. As of November 17, 2015:
http://www.rand.org/pubs/monographs/MG360.html

Davis, Paul K., Eric V. Larson, Zachary Haldeman, Mustafa Oguz, and Yashodhara Rana (2012), *Understanding and Influencing Public Support for Insurgency and Terrorism*, Santa Monica, Calif.: RAND Corporation, MG-1122-OSD. As of November 13, 2015:
http://www.rand.org/pubs/monographs/MG1122.html

Davis, Paul K., Jimmie McEver, and Barry Wilson (2002), *Measuring Interdiction Capabilities in the Presence of Anti-Access Strategies: Exploratory Analysis to Inform Adaptive Strategy for the Persian Gulf*, Santa Monica, Calif.: RAND Corporation, MR-1471-AF. As of November 17, 2015:
http://www.rand.org/pubs/monograph_reports/MR1471.html

Davis, Paul K., and Angela O'Mahony (2013), *A Computational Model of Public Support for Insurgency and Terrorism: A Prototype for More-General Social-Science Modeling*, Santa Monica, Calif.: RAND Corporation, TR-1220-OSD. As of November 13, 2015:
http://www.rand.org/pubs/technical_reports/TR1220.html

Davis, Paul K., Walter L. Perry, Ryan Andrew Brown, Douglas Yeung, Parisa Roshan, and Phoenix Voorhies (2013), *Using Behavioral Indicators to Help Detect Potential Violent Acts: A Review of the Science Base*, Santa Monica, Calif.: RAND Corporation, RR-215-NAVY. As of November 13, 2015:
http://www.rand.org/pubs/research_reports/RR215.html

Dawes, Robyn M. (1979), "The Robust Beauty of Improper Linear Models," *American Psychologist*, Vol. 34, pp. 571–582, reprinted (1982) in Daniel Kahneman, Paul Slovic, and Amos Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases*, New York: Cambridge University Press, pp. 391–407.

Duda, Richard O. (2004), *Pattern Classification 2nd Edition with Computer Manual*, Wiley-Interscience.

Duda, R. O., P. E. Hart, and N. J. Nilsson (1976), "Subjective Bayesian Methods for Rule-Based Inference Systems," *Proceedings of the June 7–10, 1976, National Computer Conference and Exposition, ACM*.

Engl, Heinz W., and Gerhard Landl (1993), "Convergence Rates for Maximum Entropy Regularization," *SIAM Journal on Numerical Analysis*, Vol. 30, No. 5, pp. 1509–1536.

Fenton, Norman, and Martin Neil (2012), *Risk Assessment and Decision Analysis with Bayesian Networks*, CRC Press.

Frontline Systems, Inc. (2015), "Standard Excel Solver—Limitations of Nonlinear Optimization." As of April 7, 2015:
http://www.solver.com/standard-excel-solver-dealing-problem-size-limits

Genest, Christian, and James V. Zidek (1986a), "Combining Probability Distributions: A Critique and an Annotated Bibliography," *Statistical Science*, Vol. 1, No. 1, pp. 114–135.

——— (1986b), "[Combining Probability Distributions: A Critique and an Annotated Bibliography]: Rejoinder," *Statistical Science*, Vol. 1, No. 1, pp. 147–148.

George, Alexander L., and Andrew Bennett (2005), *Case Studies and Theory Development in the Social Sciences*, Cambridge, Mass.: MIT Press.

Gigerenzer, Gerd, and Reinhar Selten (2002), *Bounded Rationality: The Adaptive Toolbox*, Cambridge, Mass.: MIT Press.

Goldstein, William M., and Robin M. Hogarth (1997), *Research on Judgment and Decision Making: Current Connections and Controversies*, Cambridge, UK: Cambridge University Press.

Heuer, Richards J. (1999), *Psychology of Intelligence Analysis*, Washington, D.C.: Central Intelligence Agency.

——— (2005, October 16), "How Does Analysis of Competing Hypotheses (ACH) Improve Intelligence Analysis?" As of November 16, 2015: http://www.pherson.org/wp-content/uploads/2013/06/06.-How-Does-ACH-Improve-Analysis_FINAL.pdf

Heuer, Richards J., and Randolph H. Pherson (2014), *Structured Analytic Techniques for Intelligence Analysis*, CQ Press.

Hoerl, Arthur E., and Robert W. Kennard (1970), "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, Vol. 12, No. 1, pp. 55–67.

Hogarth, Robin M. (1986), "[Combining Probability Distributions: A Critique and an Annotated Bibliography]: Comment," *Statistical Science*, Vol. 1, No. 1, pp. 145–147.

Intelligence Advanced Research Projects Activity, "Aggregative Contingent Estimation (ACE)" (undated). As of November 16, 2015: http://www.iarpa.gov/index.php/research-programs/ace

Jaynes, Edwin T. (1957a), "Information Theory and Statistical Mechanics," *Physical Review*, Vol. 106, No. 4, pp. 620–630.

——— (1957b), "Information Theory and Statistical Mechanics II," *Physical Review*, Vol. 108, No. 2, pp. 171–190.

Jaynes, Edwin T., and G. Larry Bretthorst, eds. (2003), *Probability Theory: The Logic of Science*, Cambridge, UK: Cambridge University Press.

Kahneman, Daniel (2002, December 8), *Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice*, Nobel Prize Lecture given at Aula Magna, Stockholm University. As of November 16, 2015: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2002/kahneman-lecture.html

——— (2011), *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.

Kahneman, Daniel, Paul Slovic, and Amos Tversky, eds. (1982), *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge, UK: Cambridge University Press.

Katz, Amnon (1967), *Principles of Statistical Mechanics: The Information Theoretic Approach*, San Francisco, Calif.: W. H. Freeman.

Khaleghi, Bahador, Alaa Khamis, Fakhreddine O. Karray, and Saiedeh N. Razavib (2013), "Multisensor Data Fusion: A Review of the State-of-the-Art," *Information Fusion*, Vol. 14, No. 1.

Klein, Gary, B. Moon, and R. R, Hoffman (2006a), "Making Sense of Sensemaking 1: Alternative Perspectives," *IEEE Intelligent Systems*, Vol. 2, No. 4, pp. 70–73.

——— (2006b), "Making Sense of Sensemaking 2: A Macrocognitive Model," *IEEE Intelligent Systems*, Vol. 21, No. 5, pp. 88–92.

Klein, Lawrence A. (2004), *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*, Bellingham, Wash.: SPIE.

Kotzen, Matthew (2012), "Book Review: *In Defense of Objective Bayesianism*, by Jon Williamson," *Mind*, Vol. 120, No. 480, pp. 1324–1330.

Kullback, Solomon, and Richard A. Leibler (1951), "On Information and Sufficiency," *Annals of Mathematical Statistics*, Vol. 22, No. 1, pp. 79–86.

Laskey, Kathryn (1996), "Model Uncertainty: Theory and Practical Implications," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, Vol. 26, No. 3, pp. 340–348.

Laskey, Kathryn, Ghazi Alghamdi, Xun Wang, Daniel Barbará, Tom Shackelford, Ed Wright, and Julie Fitzgerald (2004), "Detecting Threatening Behavior Using Bayesian Networks," in *Proceedings of BRIMS Conference (Behavior Representation in Modeling and Simulation)*.

Legendre, Adrien-Marie (1805), *New Methods for the Determination of the Orbits of Comets*, Paris: F. Didiot.

List, Christian, and Ben Polak (2010, March), "Introduction to Judgment Aggregation," *Journal of Economic Theory*, Vol. 145, No. 2, pp. 441–466.

List, Christian, and Clemens Puppe (2009), "Judgment Aggregation: A Survey," in Paul Anand, Prasanta Pattanaik, and Clemens Puppe, eds., *Handbook of Rational and Social Choice*, Oxford, UK: Oxford University Press.

Lempert, Robert J., David G. Groves, Steven W. Popper, and Steven C. Bankes (2006, April), "A General Analytic Method for Generating Robust Strategies and Narrative Scenarios," *Management Science*, Vol. 4, pp. 514–528.

Lumina Decision Systems (undated), "Mixture Distribution." As of November 19, 2015:
http://wiki.analytica.com/index.php?title=Mixture_distribution

——— (2015), homepage. As of November 19, 2015: http://www.lumina.com

Maki, Uskali, ed. (2012), *Judgment Aggregation: A Short Introduction*, Elseveier Science Publishers.

Mendelson, Elliot (4th ed., 1997; 1st ed., 1964), *Introduction to Mathematical Logic*, New York: Van Nostrand Reinhold.

Mohammad-Djafari, Ali, Jean Francois Giovanelli, Guy Demoment, and Jérôme Idler (2002, April), "Regularization, Maximum Entropy and Probabilistic Methods in Mass Spectrometry Data Processing Problems," *International Journal of Mass Spectrometry*, Vol. 215, Nos. 1–3, pp. 175–193.

Morgan, Patrick M. (2003), *Deterrence Now (Cambridge Studies in International Relations)*, Cambridge, UK: Cambridge University Press.

Morgan, M. Granger, and Max Henrion (1992), *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, New York: Cambridge University Press.

Morral, Andrew R., Carter C. Price, David S. Ortiz, Bradley Wilson, Tom LaTourrette, Blake W. Mobley, Shawn McKay, and Henry H. Willis (2012), *Modeling Terrorism Risk to the Air Transportation System: An Independent Assessment of TSA's Risk Management Analysis Tool and Associated Methods*, Santa Monica, Calif.: RAND Corporation, MG-1241-TSA. As of November 19, 2015: http://www.rand.org/pubs/monographs/MG1241.html

Myung, Jae, Sridhar Ramahoorti, and Andrew D. Bailey (1996), "Maximum Entropy Aggregation of Expert Predictions," *Management Science*, Vol. 42, No. 10, pp. 1420–1436.

National Commission on Terrorist Attacks (2004, July 22), *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States*, New York: W.W. Norton & Company.

National Research Council (2008), *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment*, Washington, D.C.: The National Academies Press.

——— (2014), *U.S. Air Force Strategic Deterrence Analytic Capabilities: An Assessment of Methods, Tools, and Approaches for the 21st Century Security Environment*, Washington, D.C.: National Academies Press.

O'Hagan, Anthony, and Jeremy E. Oakley (2004, July–September), "Probability Is Perfect, But We Can't Elicit It Perfectly," *Reliability Engineering & System Safety*, Vol. 85, Nos. 1–3), pp. 239–248.

Osherson, Daniel, and Moshe Y. Vardi (2006, July), "Aggregating Disparate Estimates of Chance," *Games and Economic Behavior*, Vol. 56, No. 1, pp. 148–173.

Pearl, Judea (2nd ed., 2009; 1st ed., 2001), *Causality: Models, Reasoning, and Inference*, Cambridge, Mass.: Cambridge University Press.

Pennington, Nancy, and Reid Hastie (1988, July), "Explanation-Based Decision Making: Effects of Memory Structure on Reasoning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, pp. 521–533.

Perry, Walter L. (2011), "Evidential Reasoning in Support of Counterinsurgency Intelligence Operations: Combining Evidence from Disparate Sources," in *Proceedings of 28th ISMOR, August 29–September 2*.

Pourret, Oliver, Patrick Naim, and Bruce Marcot (2008), *Bayesian Networks: Practical Guide to Applications*, Chichester, UK: Wiley.

Predd, Joel B., Daniel N. Osherson, Sanjeev R. Kulkarni, and H. Vincent Poor (2008), "Aggregating Probabilistic Forecasts from Incoherent and Abstaining Experts," *Decision Analysis*, Vol. 5, No. 4, pp. 177–189.

RAND Corporation (2015), "Robust Decision Making." As of November 16, 2015:
http://www.rand.org/topics/robust-decision-making.html

Rovner, Joshua (2011), *Fixing the Facts*, Ithaca, N.Y.: Cornell.

Satopää, Ville A., Jonathan Baron, Dean P. Foster, Barbara A. Mellers, Philip E. Tetlock, and Lyle H. Ungar (2014, April–June), "Combining Multiple Probability Predictions Using a Simple Logit Model," *International Journal of Forecasting*, Vol. 30, No. 2, pp. 344–356.

Shafer, Glenn (1976), *A Mathematical Theory of Evidence*, Princeton, N.J.: Princeton University Press.

Shafer, Glenn, and A. P. Dempster (undated), "Dempster-Shafer Theory." As of November 16, 2015:
http://www.glennshafer.com/assets/downloads/articles/article48.pdf

Shafer, Glenn, and Judea Pearl, eds. (1990), *Bayesian and Belief-Function Formalisms for Evidential Reasoning: A Conceptual Analysis*, San Mateo, Calif.: Morgan Kaufmann Publishers Inc.

Shane, Scott (2015, April 9), "Former F.B.I. Agent Sues, Claims Retaliation Over Misgivings in Anthrax Case," *New York Times*, p. A21.

Shannon, Claude (1948), "A Mathematical Theory of Communications," *Bell Systems Technical Journal*, Vol. 27, pp. 370–423; 623–656.

Simon, Herbert A. (1978), "Rational Decision-Making in Business Organizations: Nobel Prize Lecture," in Assar Lindbeck, ed., *Nobel Lectures, Economics, 1969–1980*, Singapore: World Scientific Publishing Co.

Smarandache, Florentin, and Jean Dezert, eds. (2009a), *Advances and Applications of DSmT for Information Fusion*, Rehoboth, Del.: American Research Press.

——— (2009b), "An Introduction to DSmT," in Florentin Smarandache, and Jean Dezert, eds., *Advances and Applications of DSmT for Information Fusion*, Rehoboth, Del.: American Research Press.

Sterman, John D. (2000), *Business Dynamics: Systems Thinking and Modeling for a Complex World*, Boston: McGraw-Hill.

Strom, Kevin J., John Hollywood, Mark Pope, Garth Weintraub, Crystal Daye, and Don Gemeinhardt (2010), *Building on Clues: Examining Successes and Failures in Detecting U.S. Terrorist Plots, 1999–2009*, Durham, N.C.: Institute for Homeland Security Solutions.

Thagard, Paul (2003), "Why Wasn't O.J. Convicted? Emotional Coherence in Legal Inference," *Cognition and Emotion*, Vol. 17, No. 3, pp. 361–383.

Tibshirani, Robert (1996), "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, pp. 267–288.

Williamson, Jon (2010), *In Defense of Objective Bayesianism*, Oxford, UK: Oxford University Press.

Wright, Ed, and Bob Schrag (2014), "Automated Construction of Bayesian Networks from Qualitative Knowledge," unpublished briefing, McClean, Va.: Haystax Advanced Threat Analytics.

Zou, Hui, and Trevor Hastie (2005), "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 67, No. 2, pp. 301–320.

An element of thwarting terrorist attacks is observing suspicious individuals over time with such diverse means as scanners and other devices, travel records, behavioral observations, and intelligence sources. Such observations provide data that are often both complex and "soft"—i.e., qualitative, subjective, fuzzy, or ambiguous—and also contradictory or even deceptive. Analysts face the challenge of *heterogeneous information fusion*—that is, combining such data to form a realistic assessment of threat. This report presents research on various heterogeneous information fusion methods and describes a research prototype system for fusing uncertainty-sensitive heterogeneous information. The context is counterterrorism, for both military and civilian applications, but the ideas are also applicable in intelligence and law enforcement.

NATIONAL SECURITY RESEARCH DIVISION

www.rand.org

$32.00

ISBN-10 0-8330-9277-4
ISBN-13 978-0-8330-9277-9

RR-1200-NAVY